

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 19:23:59 ; Search time 245 Seconds
(without alignments)
3996.308 Million cell updates/sec

Title: US-10-772-636-64
Perfect score: 10042
Sequence: 1 MVHVARLLLLLTLFLRTDA.....YQFSYRAALEYLGSGFDHYAT 1912

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10042	100.0	1912	1 PTPD_HUMAN	P23468 homo sapien
2	9280.5	92.4	1894	2 Q64487	Q64487 mus musculus
3	8991.5	89.5	1896	2 Q91AJ1	Q91AJ1 xenopus lae
4	7386.5	73.6	1948	1 PTNS_HUMAN	Q13332 homo sapien
5	7110	70.8	1904	2 Q64699	Q64699 mus musculus
6	7032	70.0	1897	1 PTPF_HUMAN	P10586 homo sapien
7	7004.5	69.8	1898	2 Q86W50	Q86W50 homo sapien
8	6985.5	69.6	1898	2 Q9EQ17	Q9EQ17 mus musculus
9	6981.5	69.5	1898	2 Q64604	Q64604 x protein-t
10	6981	69.5	1887	2 Q9QW67	Q9QW67 rattus sp.
11	6977.5	69.5	1863	2 Q64605	Q64605 rattus norv
12	6750.5	67.2	1788	2 Q91AJ0	Q91AJ0 xenopus lae
13	6308.5	62.8	1254	2 Q8VBV0	Q8VBV0 mus musculus
14	5871.5	58.5	1499	2 Q9UM81	Q9UM81 homo sapien
15	5794.5	57.2	1502	2 Q7T117	Q7T117 mus musculus
16	5748.5	57.0	1501	2 Q9QW00	Q9QW00 rattus sp.
17	5745.5	57.2	1501	2 Q9QW00	Q9QW00 rattus sp.
18	5640	56.2	1529	2 Q8FG86	Q8FG86 mus musculus
19	4935.5	49.1	1889	2 Q7Q0X2	Q7Q0X2 anopheles g
20	4879	48.6	2029	2 Q9V1S8	Q9V1S8 drosophila
21	4877	48.6	2029	1 LAR_DROME	P16621 drosophila
22	4594.5	45.8	1191	2 Q7Z3X4	Q7Z3X4 homo sapien
23	4127	41.1	2051	2 Q44328	Q44328 hirudo medi
24	3942	39.3	749	2 Q8R169	Q8R169 mus musculus
25	3822	38.1	730	2 Q8WX65	Q8WX65 homo sapien
26	3785	37.7	1597	2 Q960M3	Q960M3 drosophila
27	3541	35.3	857	2 Q90YJ4	Q90YJ4 brachydanio
28	3372	33.6	1437	2 Q44329	Q44329 hirudo medi
29	3359	33.4	2200	1 LAR_CAEL	Q9BMN8 caenorhabdi
30	3044	30.3	615	2 Q91A18	Q91A18 xenopus lae
31	2996	29.8	1231	2 Q17024	Q17024 anopheles g

32	2820	28.1	582	2	Q64696	Q64696 mus musculus
33	2504	24.9	508	2	Q90YJ5	Q90YJ5 brachydanio
34	2404	23.9	468	2	Q91BA2	Q91BA2 potamotrygo
35	2381	23.7	468	2	Q91BA0	Q91BA0 potamotrygo
36	2339	23.3	469	2	Q9NL11	Q9NL11 branchiost
37	2304	22.9	468	2	Q91BA5	Q91BA5 potamotrygo
38	2242	22.3	460	2	Q82917	Q82917 rattus norv
39	2226.5	22.2	469	2	Q9NL08	Q9NL08 eptatretus
40	2093	20.8	468	2	Q9NL06	Q9NL06 eptatretus
41	2045.5	20.4	1216	2	Q81120	Q81120 caenorhabdi
42	1954	19.5	398	2	Q62604	Q62604 rattus norv
43	1676	16.7	383	2	Q8MTN0	Q8MTN0 culicoides
44	1651	16.4	472	2	Q9NL02	Q9NL02 eptatretus
45	1586	15.8	793	2	Q91V35	Q91V35 m protein t

ALIGNMENTS

RESULT 1

ID	PTPD_HUMAN	STANDARD;	PRT;	1912 AA.
AC	P23468;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Protein-tyrosine phosphatase delta precursor (EC 3.1.3.48) (R-PTP-delta).			
GN	Name=PTPRD;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-1178.			
RP	MEDLINE=95204468; PubMed=7896816; DOI=10.1074/jbc.270.12.6722;			
RA	Pulido R., Krueger N.X., Serra-Pages C., Saito H., Streuli M.;			
RT	"Molecular characterization of the human transmembrane protein-tyrosine phosphatase delta. Evidence for tissue-specific expression of alternative human transmembrane protein-tyrosine phosphatase delta isoforms.";			
RT	J. Biol. Chem. 270:6722-6728(1995).			
RN	[2]			
RP	SEQUENCE OF 390-1912 FROM N.A.			
RC	TISSUE=Placenta;			
RA	MEDLINE=91006018; PubMed=2170109;			
RA	Krueger N.X., Streuli M., Saito H.;			
RT	"Structural diversity and evolution of human receptor-like protein tyrosine phosphatases.";			
RL	EMBO J. 9:3241-3252(1990).			
RN	[3]			
RP	INTERACTIONS WITH PPPIA1; PPPIA2 AND PPPIA3.			
RP	MEDLINE=98288299; PubMed=9624153; DOI=10.1074/jbc.273.25.15611;			
RA	Serra-Pages C., Medley O.G., Tang M., Hart A., Streuli M.;			
RT	"Liprins, a family of LAR transmembrane protein-tyrosine phosphatase-interacting proteins.";			
RL	J. Biol. Chem. 273:15611-15620(1998).			
RN	[4]			
RP	SPLICE ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).			
RP	PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;			
RA	Hillman R.T., Green R.E., Brenner S.E.;			
RT	"An unappreciated role for RNA surveillance.";			
RL	Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).			
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein tyrosine + phosphate.			
CC	-1- SUBUNIT: Interacts with PPPIA1, PPPIA2 and PPPIA3.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event-Alternative splicing; Named isoforms=3;			
CC	Comment=Additional isoforms seem to exist;			
CC	Name=1;			
CC	Isoid=23468-1; Sequence=Displayed;			
CC	Note=May be produced at very low levels due to a premature stop			

CC codon in the mRNA, leading to nonsense-mediated mRNA decay;
 CC Name=2; Synonym=Kidney;
 CC IsoId=E23468-2; Sequence=VSP_005147, VSP_005148, VSP_005149;
 CC Name=3; Synonym=Fetal brain;
 CC IsoId=E23468-3; Sequence=VSP_005150;
 CC -1- PPM: A cleavage occurs that separates the extracellular domain
 CC from the transmembrane segment.
 CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC Receptor class 2A subfamily.
 CC -1- SIMILARITY: Contains 8 fibronectin type III domains.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 2 protein-tyrosine phosphatase domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L38929; AAC1749.1; -;
 CC EMBL: X54133; CAA38068.1; -;
 CC PIR: A56178; A56178.
 CC HSP: P10586; ILAR.
 CC Genew: HGNC:9668; PTPRD.
 CC MIM: 601598; -;
 CC GO: GO:0005887; C: integral to plasma membrane; TAS.
 CC GO: GO:0005001; P: transmembrane receptor protein tyrosine pho. .; TAS.
 CC GO: GO:0006470; P: protein amino acid dephosphorylation; TAS.
 CC GO: GO:0007185; P: transmembrane receptor protein tyrosine pho. .; TAS.
 CC InterPro: IPR003961; FN_III.
 CC InterPro: IPR008957; FN_III-like.
 CC InterPro: IPR003962; FNIII subd.
 CC InterPro: IPR007110; Ig-like.
 CC InterPro: IPR003598; Ig_c2.
 CC InterPro: IPR000387; TYR phosphatase.
 CC InterPro: IPR000242; Tyr_PP.
 CC Pfam: PF00041; fn3; 8.
 CC Pfam: PF00047; Ig; 3.
 CC Pfam: PF00102; Y_phosphatase; 2.
 CC PRINTS: PR00014; FNTYPEIII.
 CC PRINTS: PR00700; PTPPHPTASE.
 CC SMART: SM00060; FN3; 8.
 CC SMART: SM00408; IGC2; 3.
 CC SMART: SM00194; PTPC; 2.
 CC PROSITE: PS00853; FN3; 8.
 CC PROSITE: PS00835; IG LIKE; 3.
 CC PROSITE: PS00383; TYR_PHOSPHATASE_1; 2.
 CC PROSITE: PS00056; TYR_PHOSPHATASE_2; 2.
 CC PROSITE: PS00055; TYR_PHOSPHATASE_PTP; 2.
 CC KW Alternative splicing; Glycoprotein; Hydrolase; Immunoglobulin domain;
 KW Protein phosphatase; Receptor; Repeat; Signal; Transmembrane.
 CC SIGNAL 1 20
 FT CHAIN 1 1912 Protein-tyrosine phosphatase delta.
 FT DOMAIN 21 1265 Extracellular (Potential).
 FT TRANSMEM 1266 1290 Potential.
 FT DOMAIN 1291 1912 Cytoplasmic (Potential).
 FT DOMAIN 24 114 Ig-like C2-type 1.
 FT DOMAIN 126 224 Ig-like C2-type 2.
 FT DOMAIN 236 318 Ig-like C2-type 3.
 FT DOMAIN 323 411 Fibronectin type-III 1.
 FT DOMAIN 417 511 Fibronectin type-III 2.
 FT DOMAIN 516 604 Fibronectin type-III 3.
 FT DOMAIN 609 706 Fibronectin type-III 4.
 FT DOMAIN 711 819 Fibronectin type-III 5.
 FT DOMAIN 824 913 Fibronectin type-III 6.
 FT DOMAIN 918 1013 Fibronectin type-III 7.
 FT DOMAIN 1017 1103 Fibronectin type-III 8.
 FT DOMAIN 1375 1618 Protein-tyrosine phosphatase 1.
 FT DOMAIN 1619 1912 Protein-tyrosine phosphatase 2.
 FT ACT_SITE 1553 phosphotyrosine intermediate (By
 FT similarity).

FT	ACT_SITE	1844	1844	1844	Phosphotyrosine intermediate (By similarity)
FT	SITE	1175	1178	1178	Cleavage (Potential).
FT	CARBOHYD	254	254	254	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	299	299	299	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	724	724	724	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	832	832	832	N-linked (GlcNAc. . .) (Potential).
FT	VARSPLIC	181	189	189	Missing (in isoform 2). /FTId=VSP_005147.
FT	VARSPLIC	226	229	229	Missing (in isoform 2). /FTId=VSP_005148.
FT	VARSPLIC	775	783	783	Missing (in isoform 2). /FTId=VSP_005149.
FT	VARSPLIC	609	1137	1137	Missing (in isoform 3). /FTId=VSP_005150.
FT	MUTAGEN	1178	1178	1178	R->A: 2.5-fold reduction in cleavage.
SQ	SEQUENCE	1912 AA;	214759 MW;	3A88BCD32182E26 CRC64;	
Query Match 100.0%; Score 10042; DB 1; Length 1912;					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 1912; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	MVHVARELLLLLTFFLRDAETPPRTTRTPVDQTVGGVSGVASFICQATGDPKPIVWNKK	60		
Db	1	MVHVARELLLLLTFFLRDAETPPRTTRTPVDQTVGGVSGVASFICQATGDPKPIVWNKK	60		
Qy	61	GKKVSNORFEVIEPDDGGSGVLRIQPLTRPRDEAIYECVASNNVGEISVSTRITVLREDQ	120		
Db	61	GKKVSNORFEVIEPDDGGSGVLRIQPLTRPRDEAIYECVASNNVGEISVSTRITVLREDQ	120		
Qy	121	IPRGFPIDMGPKQKVVERTRTATMLCAASGNPDPEITWPKDFLPDVTNNNGRIKQLRS	180		
Db	121	IPRGFPIDMGPKQKVVERTRTATMLCAASGNPDPEITWPKDFLPDVTNNNGRIKQLRS	180		
Qy	181	ESIGGTPIRGALQIEQSEESDQGYECVATNSAGTRYSAAPANLYVRELREVRVPPPSI	240		
Db	181	ESIGGTPIRGALQIEQSEESDQGYECVATNSAGTRYSAAPANLYVRELREVRVPPPSI	240		
Qy	241	PPTNHEIMPGGSVNITCVAVGSPMPYKWLGAEDLTPEDDMPGRNVLENDVRSQANY	300		
Db	241	PPTNHEIMPGGSVNITCVAVGSPMPYKWLGAEDLTPEDDMPGRNVLENDVRSQANY	300		
Qy	301	TCVAMSTGLVIEALQITVAKLPKPGTPVVTSTATSTILTWSGNGPEPVSYYIIQHP	360		
Db	301	TCVAMSTGLVIEALQITVAKLPKPGTPVVTSTATSTILTWSGNGPEPVSYYIIQHP	360		
Qy	361	KNSEELYKEIDGVATTRYSVAGLSPYSDYEPVVAVNNIGRGPSEPVLTTQTSQAPSSA	420		
Db	361	KNSEELYKEIDGVATTRYSVAGLSPYSDYEPVVAVNNIGRGPSEPVLTTQTSQAPSSA	420		
Qy	421	PRDVOARMLSTTILVQWKEPEENGQIQGYRVVYTMPTQHVNNMKNHVNADSOITIG	480		
Db	421	PRDVOARMLSTTILVQWKEPEENGQIQGYRVVYTMPTQHVNNMKNHVNADSOITIG	480		
Qy	481	NLVPQKTSYVKVLAFTSIGDPLSSDIQVITQTGTGVPQPLNFKAEPSESETILLSWTTPR	540		
Db	481	NLVPQKTSYVKVLAFTSIGDPLSSDIQVITQTGTGVPQPLNFKAEPSESETILLSWTTPR	540		
Qy	541	SDTIANYELVYKDGEGHEEQRIITIEPGTSYRLOGKPNLSYFRLAARSPQGLGASTAEI	600		
Db	541	SDTIANYELVYKDGEGHEEQRIITIEPGTSYRLOGKPNLSYFRLAARSPQGLGASTAEI	600		
Qy	601	SARTMQSKPSAPPODISCTSPSSSTILVSWOPPVVEKONGIITEYSIKYTAVDGDDKPH	660		
Db	601	SARTMQSKPSAPPODISCTSPSSSTILVSWOPPVVEKONGIITEYSIKYTAVDGDDKPH	660		
Qy	661	EILGIPSDTTKYLLLEQLEKTEYRITVTAHTDVGPGESLSVLIRTNEDVSPGPKVVEV	720		
Db	661	EILGIPSDTTKYLLLEQLEKTEYRITVTAHTDVGPGESLSVLIRTNEDVSPGPKVVEV	720		
Qy	721	EAVNSTSVKWSRSPVFNKQHGQIRGQVHVYVRMENGEPKQPMKQVMLADAQWEPDDT	780		
Db	721	EAVNSTSVKWSRSPVFNKQHGQIRGQVHVYVRMENGEPKQPMKQVMLADAQWEPDDT	780		

QY 781 TEHDMIIISGLQETSYSLSLTVAITTKGDSKSKPKLVSTTGAVPGKPRLVINHTQNTAL 840
Db 781 TEHDMIIISGLQETSYSLSLTVAITTKGDSKSKPKLVSTTGAVPGKPRLVINHTQNTAL 840
QY 841 IQWHPPVDTFGLQGVRLKFGKMEPLTTLFSEKEDHFTATDTHKGSYVPRLSARNK 900
Db 841 IQWHPPVDTFGLQGVRLKFGKMEPLTTLFSEKEDHFTATDTHKGSYVPRLSARNK 900
QY 901 VGFGEWVKBIISPEEPTGTFPQNLHSEGTSTSVQLSWQPPVLAERNGIITKYLLYRD 960
Db 901 VGFGEWVKBIISPEEPTGTFPQNLHSEGTSTSVQLSWQPPVLAERNGIITKYLLYRD 960
QY 961 INIPLLPHEQLIVPADTMTLTLGLKPDYTDYKVRHAHSGKPGTSPSVQFTPLVDQVF 1020
Db 961 INIPLLPHEQLIVPADTMTLTLGLKPDYTDYKVRHAHSGKPGTSPSVQFTPLVDQVF 1020
QY 1021 AKNFHVKAAMKTSVLLSWEIPEYNYSAMPFKLYDGMKVEVDGRATQKLVNLKPKES 1080
Db 1021 AKNFHVKAAMKTSVLLSWEIPEYNYSAMPFKLYDGMKVEVDGRATQKLVNLKPKES 1080
QY 1081 YSFVLTNRGNSAGGLQHRVTAKTAPDVLRTKPAFIGNLDMITVQLPEVPANENIKGY 1140
Db 1081 YSFVLTNRGNSAGGLQHRVTAKTAPDVLRTKPAFIGNLDMITVQLPEVPANENIKGY 1140
QY 1141 YIIIVPLKSRGKFKIPWESPEMELDELLKESIRKRSIRYGREVELKPYIAAHFDVLP 1200
Db 1141 YIIIVPLKSRGKFKIPWESPEMELDELLKESIRKRSIRYGREVELKPYIAAHFDVLP 1200
QY 1201 TEFTLGDHKGFTNKLOSGQEVVFFVLAHMEHAEKMYATSPYDPVMSMDLDPOPI 1260
Db 1201 TEFTLGDHKGFTNKLOSGQEVVFFVLAHMEHAEKMYATSPYDPVMSMDLDPOPI 1260
QY 1261 TDEEGLIWWGVPLAVVPIICIVAILLYKRAESDSRKSSIPNNKEIPSHHPTDVE 1320
Db 1261 TDEEGLIWWGVPLAVVPIICIVAILLYKRAESDSRKSSIPNNKEIPSHHPTDVE 1320
QY 1321 LRLNFQTPGASHHPPIILELAHIERLKANDNLKFSQEVESIDPGQQTWEHNSLEVN 1380
Db 1321 LRLNFQTPGASHHPPIILELAHIERLKANDNLKFSQEVESIDPGQQTWEHNSLEVN 1380
QY 1381 KPKRYANVIAVDHRSVLLSAIEGIPGSDYVNNYIDGVRKONAVIATQSGIPETFGDFW 1440
Db 1381 KPKRYANVIAVDHRSVLLSAIEGIPGSDYVNNYIDGVRKONAVIATQSGIPETFGDFW 1440
QY 1441 RMIEWQRSATVVMTKLEERSRVKCDQWPSRGTEHGLVQVTLTDLTVELATYCVRTAL 1500
Db 1441 RMIEWQRSATVVMTKLEERSRVKCDQWPSRGTEHGLVQVTLTDLTVELATYCVRTAL 1500
QY 1501 YKGSSEKREVRQFTAMPDGHGVEHPTPLFLRVRKTCNPPDAGPMVHCSAGVRT 1560
Db 1501 YKGSSEKREVRQFTAMPDGHGVEHPTPLFLRVRKTCNPPDAGPMVHCSAGVRT 1560
QY 1561 GCFVIDAMLERIKHEKTVDIYGHVTLMAORNYMVQTEDOYIFTHDALLAEAVTCNGTEV 1620
Db 1561 GCFVIDAMLERIKHEKTVDIYGHVTLMAORNYMVQTEDOYIFTHDALLAEAVTCNGTEV 1620
QY 1621 PARNLYAYIKLTOIETGENTVGMLEFKRLASSKAHSTRFISANLPCNKFKNRLVNIIMP 1680
Db 1621 PARNLYAYIKLTOIETGENTVGMLEFKRLASSKAHSTRFISANLPCNKFKNRLVNIIMP 1680
QY 1681 YESTRVCLQIRGVEGSDYINASFIDGVRQKAYIATQGLAETTEDFWMLWEHNSITV 1740
Db 1681 YESTRVCLQIRGVEGSDYINASFIDGVRQKAYIATQGLAETTEDFWMLWEHNSITV 1740
QY 1741 VMLTKLRMGREKCHOYPAERSARYQFVVDPMAYENMPQYILREFKVTVDARQOSRTV 1800
Db 1741 VMLTKLRMGREKCHOYPAERSARYQFVVDPMAYENMPQYILREFKVTVDARQOSRTV 1800
QY 1801 RQFQFTDWPQGVKPSGSGFIDFGQVHKTKEQFGDGPISVHCSAGVGRGTGVTITLSIV 1860
Db 1801 RQFQFTDWPQGVKPSGSGFIDFGQVHKTKEQFGDGPISVHCSAGVGRGTGVTITLSIV 1860

QY 1861 LERMYEGVVDIFQTVKMLRTQRPAMVQTEDOYQFSYRAALEYLGSDHYAT 1912
Db 1861 LERMYEGVVDIFQTVKMLRTQRPAMVQTEDOYQFSYRAALEYLGSDHYAT 1912
RESULT 2
Q64487 ID Q64487 PRELIMINARY; PRT; 1894 AA.
AC Q64487; Q64488; Q64489; Q64490; Q64491; Q64492; Q64493; Q64494; Q64495; Q64496; Q64497; Q64498; Q64499; Q64500; Q64501; Q64502; Q64503; Q64504; Q64505; Q64506; Q64507; Q64508; Q64509; Q64510; Q64511; Q64512; Q64513; Q64514; Q64515; Q64516; Q64517; Q64518; Q64519; Q64520; Q64521; Q64522; Q64523; Q64524; Q64525; Q64526; Q64527; Q64528; Q64529; Q64530; Q64531; Q64532; Q64533; Q64534; Q64535; Q64536; Q64537; Q64538; Q64539; Q64540; Q64541; Q64542; Q64543; Q64544; Q64545; Q64546; Q64547; Q64548; Q64549; Q64550; Q64551; Q64552; Q64553; Q64554; Q64555; Q64556; Q64557; Q64558; Q64559; Q64560; Q64561; Q64562; Q64563; Q64564; Q64565; Q64566; Q64567; Q64568; Q64569; Q64570; Q64571; Q64572; Q64573; Q64574; Q64575; Q64576; Q64577; Q64578; Q64579; Q64580; Q64581; Q64582; Q64583; Q64584; Q64585; Q64586; Q64587; Q64588; Q64589; Q64590; Q64591; Q64592; Q64593; Q64594; Q64595; Q64596; Q64597; Q64598; Q64599; Q64600; Q64601; Q64602; Q64603; Q64604; Q64605; Q64606; Q64607; Q64608; Q64609; Q64610; Q64611; Q64612; Q64613; Q64614; Q64615; Q64616; Q64617; Q64618; Q64619; Q64620; Q64621; Q64622; Q64623; Q64624; Q64625; Q64626; Q64627; Q64628; Q64629; Q64630; Q64631; Q64632; Q64633; Q64634; Q64635; Q64636; Q64637; Q64638; Q64639; Q64640; Q64641; Q64642; Q64643; Q64644; Q64645; Q64646; Q64647; Q64648; Q64649; Q64650; Q64651; Q64652; Q64653; Q64654; Q64655; Q64656; Q64657; Q64658; Q64659; Q64660; Q64661; Q64662; Q64663; Q64664; Q64665; Q64666; Q64667; Q64668; Q64669; Q64670; Q64671; Q64672; Q64673; Q64674; Q64675; Q64676; Q64677; Q64678; Q64679; Q64680; Q64681; Q64682; Q64683; Q64684; Q64685; Q64686; Q64687; Q64688; Q64689; Q64690; Q64691; Q64692; Q64693; Q64694; Q64695; Q64696; Q64697; Q64698; Q64699; Q64700; Q64701; Q64702; Q64703; Q64704; Q64705; Q64706; Q64707; Q64708; Q64709; Q64710; Q64711; Q64712; Q64713; Q64714; Q64715; Q64716; Q64717; Q64718; Q64719; Q64720; Q64721; Q64722; Q64723; Q64724; Q64725; Q64726; Q64727; Q64728; Q64729; Q64730; Q64731; Q64732; Q64733; Q64734; Q64735; Q64736; Q64737; Q64738; Q64739; Q64740; Q64741; Q64742; Q64743; Q64744; Q64745; Q64746; Q64747; Q64748; Q64749; Q64750; Q64751; Q64752; Q64753; Q64754; Q64755; Q64756; Q64757; Q64758; Q64759; Q64760; Q64761; Q64762; Q64763; Q64764; Q64765; Q64766; Q64767; Q64768; Q64769; Q64770; Q64771; Q64772; Q64773; Q64774; Q64775; Q64776; Q64777; Q64778; Q64779; Q64780; Q64781; Q64782; Q64783; Q64784; Q64785; Q64786; Q64787; Q64788; Q64789; Q64790; Q64791; Q64792; Q64793; Q64794; Q64795; Q64796; Q64797; Q64798; Q64799; Q64800; Q64801; Q64802; Q64803; Q64804; Q64805; Q64806; Q64807; Q64808; Q64809; Q64810; Q64811; Q64812; Q64813; Q64814; Q64815; Q64816; Q64817; Q64818; Q64819; Q64820; Q64821; Q64822; Q64823; Q64824; Q64825; Q64826; Q64827; Q64828; Q64829; Q64830; Q64831; Q64832; Q64833; Q64834; Q64835; Q64836; Q64837; Q64838; Q64839; Q64840; Q64841; Q64842; Q64843; Q64844; Q64845; Q64846; Q64847; Q64848; Q64849; Q64850; Q64851; Q64852; Q64853; Q64854; Q64855; Q64856; Q64857; Q64858; Q64859; Q64860; Q64861; Q64862; Q64863; Q64864; Q64865; Q64866; Q64867; Q64868; Q64869; Q64870; Q64871; Q64872; Q64873; Q64874; Q64875; Q64876; Q64877; Q64878; Q64879; Q64880; Q64881; Q64882; Q64883; Q64884; Q64885; Q64886; Q64887; Q64888; Q64889; Q64890; Q64891; Q64892; Q64893; Q64894; Q64895; Q64896; Q64897; Q64898; Q64899; Q64900; Q64901; Q64902; Q64903; Q64904; Q64905; Q64906; Q64907; Q64908; Q64909; Q64910; Q64911; Q64912; Q64913; Q64914; Q64915; Q64916; Q64917; Q64918; Q64919; Q64920; Q64921; Q64922; Q64923; Q64924; Q64925; Q64926; Q64927; Q64928; Q64929; Q64930; Q64931; Q64932; Q64933; Q64934; Q64935; Q64936; Q64937; Q64938; Q64939; Q64940; Q64941; Q64942; Q64943; Q64944; Q64945; Q64946; Q64947; Q64948; Q64949; Q64950; Q64951; Q64952; Q64953; Q64954; Q64955; Q64956; Q64957; Q64958; Q64959; Q64960; Q64961; Q64962; Q64963; Q64964; Q64965; Q64966; Q64967; Q64968; Q64969; Q64970; Q64971; Q64972; Q64973; Q64974; Q64975; Q64976; Q64977; Q64978; Q64979; Q64980; Q64981; Q64982; Q64983; Q64984; Q64985; Q64986; Q64987; Q64988; Q64989; Q64990; Q64991; Q64992; Q64993; Q64994; Q64995; Q64996; Q64997; Q64998; Q64999; Q65000; Q65001; Q65002; Q65003; Q65004; Q65005; Q65006; Q65007; Q65008; Q65009; Q65010; Q65011; Q65012; Q65013; Q65014; Q65015; Q65016; Q65017; Q65018; Q65019; Q65020; Q65021; Q65022; Q65023; Q65024; Q65025; Q65026; Q65027; Q65028; Q65029; Q65030; Q65031; Q65032; Q65033; Q65034; Q65035; Q65036; Q65037; Q65038; Q65039; Q65040; Q65041; Q65042; Q65043; Q65044; Q65045; Q65046; Q65047; Q65048; Q65049; Q65050; Q65051; Q65052; Q65053; Q65054; Q65055; Q65056; Q65057; Q65058; Q65059; Q65060; Q65061; Q65062; Q65063; Q65064; Q65065; Q65066; Q65067; Q65068; Q65069; Q65070; Q65071; Q65072; Q65073; Q65074; Q65075; Q65076; Q65077; Q65078; Q65079; Q65080; Q65081; Q65082; Q65083; Q65084; Q65085; Q65086; Q65087; Q65088; Q65089; Q65090; Q65091; Q65092; Q65093; Q65094; Q65095; Q65096; Q65097; Q65098; Q65099; Q65100; Q65101; Q65102; Q65103; Q65104; Q65105; Q65106; Q65107; Q65108; Q65109; Q65110; Q65111; Q65112; Q65113; Q65114; Q65115; Q65116; Q65117; Q65118; Q65119; Q65120; Q65121; Q65122; Q65123; Q65124; Q65125; Q65126; Q65127; Q65128; Q65129; Q65130; Q65131; Q65132; Q65133; Q65134; Q65135; Q65136; Q65137; Q65138; Q65139; Q65140; Q65141; Q65142; Q65143; Q65144; Q65145; Q65146; Q65147; Q65148; Q65149; Q65150; Q65151; Q65152; Q65153; Q65154; Q65155; Q65156; Q65157; Q65158; Q65159; Q65160; Q65161; Q65162; Q65163; Q65164; Q65165; Q65166; Q65167; Q65168; Q65169; Q65170; Q65171; Q65172; Q65173; Q65174; Q65175; Q65176; Q65177; Q65178; Q65179; Q65180; Q65181; Q65182; Q65183; Q65184; Q65185; Q65186; Q65187; Q65188; Q65189; Q65190; Q65191; Q65192; Q65193; Q65194; Q65195; Q65196; Q65197; Q65198; Q65199; Q65200; Q65201; Q65202; Q65203; Q65204; Q65205; Q65206; Q65207; Q65208; Q65209; Q65210; Q65211; Q65212; Q65213; Q65214; Q65215; Q65216; Q65217; Q65218; Q65219; Q65220; Q65221; Q65222; Q65223; Q65224; Q65225; Q65226; Q65227; Q65228; Q65229; Q65230; Q65231; Q65232; Q65233; Q65234; Q65235; Q65236; Q65237; Q65238; Q65239; Q65240; Q65241; Q65242; Q65243; Q65244; Q65245; Q65246; Q65247; Q65248; Q65249; Q65250; Q65251; Q65252; Q65253; Q65254; Q65255; Q65256; Q65257; Q65258; Q65259; Q65260; Q65261; Q65262; Q65263; Q65264; Q65265; Q65266; Q65267; Q65268; Q65269; Q65270; Q65271; Q65272; Q65273; Q65274; Q65275; Q65276; Q65277; Q65278; Q65279; Q65280; Q65281; Q65282; Q65283; Q65284; Q65285; Q65286; Q65287; Q65288; Q65289; Q65290; Q65291; Q65292; Q65293; Q65294; Q65295; Q65296; Q65297; Q65298; Q65299; Q65300; Q65301; Q65302; Q65303; Q65304; Q65305; Q65306; Q65307; Q65308; Q65309; Q65310; Q65311; Q65312; Q65313; Q65314; Q65315; Q65316; Q65317; Q65318; Q65319; Q65320; Q65321; Q65322; Q65323; Q65324; Q65325; Q65326; Q65327; Q65328; Q65329; Q65330; Q65331; Q65332; Q65333; Q65334; Q65335; Q65336; Q65337; Q65338; Q65339; Q65340; Q65341; Q65342; Q65343; Q65344; Q65345; Q65346; Q65347; Q65348; Q65349; Q65350; Q65351; Q65352; Q65353; Q65354; Q65355; Q65356; Q65357; Q65358; Q65359; Q65360; Q65361; Q65362; Q65363; Q65364; Q65365; Q65366; Q65367; Q65368; Q65369; Q65370; Q65371; Q65372; Q65373; Q65374; Q65375; Q65376; Q65377; Q65378; Q65379; Q65380; Q65381; Q65382; Q65383; Q65384; Q65385; Q65386; Q65387; Q65388; Q65389; Q65390; Q65391; Q65392; Q65393; Q65394; Q65395; Q65396; Q65397; Q65398; Q65399; Q65400; Q65401; Q65402; Q65403; Q65404; Q65405; Q65406; Q65407; Q65408; Q65409; Q65410; Q65411; Q65412; Q65413; Q65414; Q65415; Q65416; Q65417; Q65418; Q65419; Q65420; Q65421; Q65422; Q65423; Q65424; Q65425; Q65426; Q65427; Q65428; Q65429; Q65430; Q65431; Q65432; Q65433; Q65434; Q65435; Q65436; Q65437; Q65438; Q65439; Q65440; Q65441; Q65442; Q65443; Q65444; Q65445; Q65446; Q65447; Q65448; Q65449; Q65450; Q65451; Q65452; Q65453; Q65454; Q65455; Q65456; Q65457; Q65458; Q65459; Q65460; Q65461; Q65462; Q65463; Q65464; Q65465; Q65466; Q65467; Q65468; Q65469; Q65470; Q65471; Q65472; Q65473; Q65474; Q65475; Q65476; Q65477; Q65478; Q65479; Q65480; Q65481; Q65482; Q65483; Q65484; Q65485; Q65486; Q65487; Q65488; Q65489; Q65490; Q65491; Q65492; Q65493; Q65494; Q65495; Q65496; Q65497; Q65498; Q65499; Q65500; Q65501; Q65502; Q65503; Q65504; Q65505; Q65506; Q65507; Q65508; Q65509; Q65510; Q65511; Q65512; Q65513; Q65514; Q65515; Q65516; Q65517; Q65518; Q65519; Q65520; Q65521; Q65522; Q65523; Q65524; Q65525; Q65526; Q65527; Q65528; Q65529; Q65530; Q65531; Q65532; Q65533; Q65534; Q65535; Q65536; Q65537; Q65538; Q65539; Q65540; Q65541; Q65542; Q65543; Q65544; Q65545; Q65546; Q65547; Q65548; Q65549; Q65550; Q65551; Q65552; Q65553; Q65554; Q65555; Q65556; Q65557; Q65558; Q65559; Q65560; Q65561; Q65562; Q65563; Q65564; Q65565; Q65566; Q65567; Q65568; Q65569; Q65570; Q65571; Q65572; Q65573; Q65574; Q65575; Q65576; Q65577; Q65578; Q65579; Q65580; Q65581; Q65582; Q65583; Q65584; Q65585; Q65586; Q65587; Q65588; Q65589; Q65590; Q65591; Q65592; Q65593; Q65594; Q65595; Q65596; Q65597; Q65598; Q65599; Q65600; Q65601; Q65602; Q65603; Q65604; Q65605; Q65606; Q65607; Q65608; Q65609; Q65610; Q65611; Q65612; Q65613; Q65614; Q65615; Q65616; Q65617; Q65618; Q65619; Q65620; Q65621; Q65622; Q65623; Q65624; Q65625; Q65626; Q65627; Q65628; Q65629; Q65630; Q65631; Q65632; Q65633; Q65634; Q65635; Q65636; Q65637; Q65638; Q65639; Q65640; Q65641; Q65642; Q65643; Q65644; Q65645; Q65646; Q65647; Q65648; Q65649; Q65650; Q65651; Q65652; Q65653; Q65654; Q65655; Q65656; Q65657; Q65658; Q65659; Q65660; Q65661; Q65662; Q65663; Q65664; Q65665; Q65666; Q65667; Q65668; Q65669; Q65670; Q65671; Q65672; Q65673; Q65674; Q65675; Q65676; Q65677; Q65678; Q65679; Q65680; Q65681; Q65682; Q65683; Q65684; Q65685; Q65686; Q65687; Q65688; Q65689; Q65690; Q65691; Q65692; Q65693; Q65694; Q65695; Q65696; Q65697; Q65698; Q65699; Q65700; Q65701; Q65702; Q65703; Q65704; Q65705; Q65706; Q65707; Q65708; Q65709; Q65710; Q65711; Q65712; Q65713; Q65714; Q65715; Q65716; Q65717; Q65718; Q65719; Q65720; Q65721; Q65722; Q65723; Q65724; Q65725; Q65726; Q65727; Q65728; Q65729; Q65730; Q65731; Q65732; Q65733; Q65734; Q65735; Q65736; Q65737; Q65738; Q65739; Q65740; Q65741; Q65742; Q65743; Q65744; Q65745; Q65746; Q65747; Q65748; Q65749; Q65750; Q65751; Q65752; Q65753; Q65754; Q65755; Q65756; Q65757; Q65758; Q65759; Q65760; Q65761; Q65762; Q65763; Q65764; Q65765; Q65766; Q65767; Q65768; Q65769; Q65770; Q65771; Q65772; Q65773; Q65774; Q65775; Q65776; Q65777; Q65778; Q65779; Q65780; Q65781; Q65782; Q65783; Q65784; Q65785; Q65786; Q65787; Q65788; Q65789; Q65790; Q65791; Q65792; Q65793; Q65794; Q65795; Q65796; Q65797; Q65798; Q65799; Q65800; Q65801; Q65802; Q65803; Q65804; Q65805; Q65806; Q65807; Q65808; Q65809; Q65810; Q65811; Q65812; Q65813; Q65814; Q65815; Q65816; Q65817; Q65818; Q65819; Q65820; Q65821; Q65822; Q65823; Q65824; Q65825; Q65826; Q65827; Q65828; Q65829; Q65830; Q65831; Q65832; Q65833; Q65834; Q65835; Q65836; Q65837; Q65838; Q65839; Q65840; Q65841; Q65842; Q65843; Q65844; Q65845; Q65846; Q65847; Q65848; Q65849; Q65850; Q65851; Q65852; Q65853; Q65854; Q65855; Q65856; Q65857; Q65858; Q65859; Q65860; Q65861; Q65862; Q65863; Q65864; Q65865; Q65866; Q65867; Q65868; Q65869; Q65870; Q65871; Q65872; Q65873; Q65874; Q65875; Q65876; Q65877; Q65878; Q65879; Q65880; Q65881; Q65882; Q65883; Q65884; Q65885; Q65886; Q65887; Q65888; Q65889; Q65890; Q65891; Q65892; Q65893; Q65894; Q65895; Q65896; Q65897; Q65898; Q65899; Q65900; Q65901; Q65902; Q65903; Q65904; Q65905; Q65906; Q65907; Q65908; Q65909; Q65910; Q65911; Q65912; Q65913; Q65914; Q65915; Q65916; Q65917; Q65918; Q65919; Q65920; Q65921; Q65922; Q65923; Q65924; Q65925; Q65926; Q65927; Q65928; Q65929; Q65930; Q65931; Q65932; Q65933; Q65934; Q65935; Q65936; Q65937; Q65938; Q65939; Q65940; Q65941; Q65942; Q65943; Q65944; Q65945; Q65946; Q65947; Q65948; Q65949; Q65950; Q65951; Q65952; Q65953; Q65954; Q65955; Q65956; Q65957; Q65958; Q65959; Q65960; Q65961; Q65962; Q65963; Q65964; Q65965; Q65966; Q65967; Q65968; Q65969; Q65970; Q65971; Q65972; Q65973; Q65974; Q65975; Q6597

QY 1319 VELRLNFTQPCWASHPPPILELADHIERLKANDNLKFSQYSEIDPQOQTWEHSNLE 1378
 Db 1304 VELRLNFTQPCWASHPPPILELADHIERLKANDNLKFSQYSEIDPQOQTWEHSNLE 1363
 QY 1379 VNKPNRANVAYDHSRVLSSAIEGIPGSDVYVANYIDGYRKQAYATQGSLETETGD 1438
 Db 1364 VNKPNRANVAYDHSRVLSSAIEGIPGSDVYVANYIDGYRKQAYATQGSLETETGD 1423
 QY 1439 FWRMIWORSATVVMWTKLEERSVKCDQYWPSCRTGTHGLVQVTLTDLTVELATYCVRTF 1498
 Db 1424 FWRMIWEO - EATVVMWTKLEERSVKCDQYWPSCRTGTHGLVQVTLTDLTVELHILCPDIC 1482
 QY 1499 ALYKNGSEKKEVRFOFTAMPDHGVPEHPTFLAFIRVKTCNPDAGPMVHCSAGVG 1558
 Db 1483 TL - NNGSSEKRVQFOTAMPDHGVPEHPTFLAFIRVKTCNPDAGPMVHCSAGVG 1541
 QY 1559 RTGCFIVDAMLERIKHEKTVDIYGHVTLMAQRNMYQTEQYIFIHDALEAVTCGNT 1618
 Db 1542 RTGCFIVDAMLERIKHEKTVDIYGHVTLMAQRNMYQTEQYIFIHDALEAVTCGNT 1601
 QY 1619 EVPARNLVAYIQKLTQIETGENVTGMELEFKRLASSKAHTSRFISANLPCNKFKNRLVNI 1678
 Db 1602 EVPARNLVAYIQKLTQIETGENVTGMELEFKRLASSKAHTSRFISANLPCNKFKNRLVNI 1661
 QY 1679 MPYESTRCLQPIRGVSGSDYINASFIDGYRQOKAYIATQGLAETTEDFWRMLWEHNS 1738
 Db 1662 MPYESTRCLQPIRGVSGSDYINASFIDGYRQOKAYIATQGLAETTEDFWRMLWEHNS 1721
 QY 1739 IVWMLTKLRMGREKCHQYPAERSARYQYFVDPMAEYNNPQYTLREFFKVTDARDG-QS 1797
 Db 1722 IVWMLTKLRMGREKCHQYPAERSARYQYFVDPMAEYNNPQYTLREFFKVTDARDG-QS 1779
 QY 1798 RTVRFQFTDMPQVPSGSGFIDFGVHKTKEQFGDQPSVHCSAGVGRGTGVFTL 1857
 Db 1780 RTVRFQFTDMPQVPSGSGFIDFGVHKTKEQFGDQPSVHCSAGVGRGTGVFTL 1839
 QY 1858 SIVLERMEYGVVDIFQTVKMLRTQRPANWQTEQYQFSYAALEYLGSFDHYAT 1912
 Db 1840 SIVLERMEYGVVDIFQTVKMLRTQRPANWQTEQYQFSYAALEYLGSFDHYAT 1894
 RESULT 3
 Q9IAJ1 PRELIMINARY; PRT; 1896 AA.
 ID Q9IAJ1
 AC Q9IAJ1
 DT 01-OCT-2000 (TremBLrel. 15, Created)
 DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Receptor protein tyrosine phosphatase delta.
 GN Name=XPTP-D;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20193505; PubMed=10727868; DOI=10.1016/S0925-4773(99)00345-7;
 RA Johnson K.G., Holt C.E.;
 RT "Expression of Crp-alpha, Lar, PTP-delta, and PTP-rho in the
 RL developing Xenopus visual system.";
 RL Mech. Dev. 92:291-294(2000).
 DR EMBL; AF197944; AAF43605.1; -.
 DR HSSP; P10586; ILAR.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0006470; F:protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR003962; FNIII subd.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR008957; FN III-like.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_C2.

DR InterPro; IPR000387; TYR_phosphatase.
 DR InterPro; IPR000242; Tyr_PP.
 DR Pfam; PF00041; fn3; 8.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00102; Y_phosphatase; 2.
 DR PRINTS; PR00014; PNTYPEIII.
 DR PRINTS; PR00700; PRTYPHPTASE.
 DR SMART; SM00060; FN3; 8.
 DR SMART; SM00408; IGC2; 3.
 DR SMART; SM00194; PTPC; 2.
 DR PROSITE; PS00853; FN3; 8.
 DR PROSITE; PS00835; IG_LIKE; 3.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 2.
 KW Hydrolase; Receptor.
 SQ SEQUENCE 1896 AA; 213067 MW; 08AC9003034199A4 CRC64;
 Query Match 89.5%; Score 8991.5; DB 2; Length 1896;
 Best Local Similarity 88.9%; Pred. No. 0;
 Matches 1699; Conservative 101; Mismatches 88; Indels 23; Gaps 4;
 QY 2 VHVARLLLLLTFLRTDAETPPRFRTRTPVDTGVGGVAFICATGDPKPKIVNNKG 61
 Db 9 MNIARPVVLLCFLHAGAETPPKLTPTVDQIGVGGVAFICATGDPKPKIVNNKG 68
 QY 62 KKVSNQRFEVIEFDGSGSVLRIQPLRTPRDEAIECVASNNVGEISVSTRITVLREDO 121
 Db 69 KKVSNQRFEVIEFDGSGSVLRIQPLRTPRDEAIECVASNNVGEISVSTRITVLREDO 128
 QY 122 PRGPTTDMGQPKVVERTRTATMLCAASGNDPDEITWFKDFLPVDTNNNGRIKQLRSE 181
 Db 129 PRGPTTDMGQPKVVERTRTATMLCAASGNDPDEITWFKDFLPVDTNNNGRIKQLRSE 187
 QY 182 SIGGTPIRGALQIQSSESDQKYECAVATNSAGTRYSAPANLYVRELEVRVRPPRESIP 241
 Db 188 -----GALQIQSSESDQKYECAVATNSAGTRYSAPANLYVRELEVRVRPPRESIP 235
 QY 242 PTNHEIMPGGSVNITCVAVGSPMPYVVKMLGAEDLTPEDDMPIGRNVLELNDVROSANYT 301
 Db 236 PTNHEIMPGGSVNITCVAVGSPMPYVVKMLGAEDLTPEDDMPIGRNVLELNDVROSANYT 295
 QY 302 CVANSTGLVIEAIAQITVKALPKPPTPVVTESTATSTLTWDSGNPFPVYIIQHKPK 361
 Db 296 CVANSTGLVIEAIAQINVKALPKPPTPVVTESTATSTLTWDSGNPFPVYIIQHKPK 355
 QY 362 NSEELYKEIDGVAATRYSVAGLSYDVEPRVAVANNIGRPPSEPVLTOTSEQAPSSAP 421
 Db 356 SSEEQYKEIDGVAATRYSVAGLSYDVEPRVAVANNIGRPPSEPVLTOTSEQAPSSAP 415
 QY 422 RDVQARMLSTTILVQWKEPEEPNGQIQGYRVYVYTMPTQHVNNMMKHNVDASQITIGN 481
 Db 416 RNQVQARMLSTTILVQWKEPEEPNGQIQGYRVYVYTMPTQHVNNMMKHNVDASQITIGN 475
 QY 482 LVPOKTSYVKVLAFTSISDGLPSSDIQVITQTVGQPLNFKAPESESTILLSWTPPRS 541
 Db 476 LEPOKTSYVKVLAFTSISDGLPSSDIQVITQTVGQPLNFKAPESESTILLSWTPPRS 535
 QY 542 DTIANVELVYKDGEGHEEORITTEBPTSRYLQGLKPNLSYVFRLAARSPOGLGASTAIS 601
 Db 536 DTISSYDLVYKDGHEAEV-ITIDPATSYRLQGLKPNLSYVFRLAARSPOGLGASTAIS 594
 QY 602 ARTMQSKSPAPPQDISCTSPSSSTILVSWQPPVPEKQNGIITEYSIKYTAVDGDDKPE 661
 Db 595 ARTMQSKSPAPPQDIRCNSQSSTILVSWQPPVPEKQNGIITEYSIKYTAVDGDDKPE 654
 QY 662 ILGTPSDTTKYLLEQLEKWTYRTVTAHTDVGCPESLSVLIRTNEDVPSPKPRKVEVE 721
 Db 655 ILGTPSDTTKYLLEQLEKWTYRTVTAHTDVGCPESLSVLIRTNEDVPSPKPRKVEVE 714
 QY 722 AVNSTSVKWSRSPVPMKQHQIRGYVHYVYRMENGPCKGPMKDVMLADAEQWEDDTT 781
 Db 715 AVNSTSVKWSRSPVPMKQHQIRGYVHYVYRMENGPCKGPMKDVMLADAEQWEDDTT 767

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2005, 19:59:42 ; Search time 2966 Seconds
(without alignments)
12500.115 Million cell updates/sec

Title: US-10-772-636-63
Perfect score: 6263
Sequence: 1 gctaactcaaggagacgtc.....aagtgaatttaaaaaagtt 6263

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2359870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6263	100.0	6263	11	ADI32147 Human CDN
2	6263	100.0	6263	13	ADR40182 Human pro
3	2473	39.5	6363	10	ADD18743 Human dis
4	2473	39.5	6363	11	ADI31553 Human CDN
5	2398	38.3	6000	2	AAQ86478 Human pTP
6	2398	38.3	6000	2	AAT85389 Human pro
7	2398	38.3	6000	2	AAX06095 Human pro
8	2398	38.3	6000	4	AAS02144 Human pro
9	2374	37.9	5913	12	ADQ84008 Human tum
10	2365.4	37.8	4555	2	AAQ94311 Tyrosine
11	2359	37.7	7741	4	AAS22448 Human CDN
12	2357.4	37.6	7702	3	AAZ91908 Human pro
13	2357.4	37.6	7702	3	AAZ59133 LAR tyros
14	2357.4	37.6	7702	3	AAA88739 Human pro
15	2357.4	37.6	7702	10	ADD18741 Human dis
16	2357.4	37.6	7702	10	ADK61221 Ovarian c
17	2357.4	37.6	7702	11	ADI32010 Human CDN
18	2357.4	37.6	7702	12	ADP18653 Human TAT
19	2357.4	37.6	7702	13	ADJ33475 Human leu
20	2357.4	37.6	7724	13	ADJ33489 Human LAR

21	2357.4	37.6	7945	5	ABV27897 Human pro
c 22	2348	37.5	7705	4	AAH98405 Human EST
c 23	2348	37.5	7705	4	AAS22684 Human CDN
24	2346.6	37.5	6545	10	ADB52922 Primary r
25	2307.8	36.8	6734	6	ABI99344 Mouse iac
26	2303.2	36.8	8232	11	ACN89826 Breast ca
27	2233.6	35.7	7007	12	ADQ85067 Human tum
28	2173.8	34.7	6469	10	ADB58330 Toxicity r
29	2173.8	34.7	6469	10	ADB52889 Primary r
30	1901.8	30.4	6050	9	AAD57343 Human kin
31	1517.6	24.2	4836	12	ADQ63853 Novel hum
32	1345.8	21.5	4078	10	ACA56534 Human sig
33	1345.8	21.5	4078	12	ADI56330 Human pol
34	1300	20.8	5690	2	AAQ86902 Rat recep
35	1290	20.6	6143	13	ADJ33482 Mouse leu
36	1245.2	19.9	5412	6	ABI99920 Rat mucoc
37	1200	19.2	4600	13	ADR06902 Full leng
38	1196.6	19.1	4190	10	ADB58003 Toxicity-
39	1095.8	17.5	3467	3	AZ91907 Human pro
40	1095.8	17.5	3467	3	AAZ59132 LAR tyros
41	953	15.2	1200	2	AAQ94313 Tyrosine
42	851.2	13.6	7630	4	ABL30223 Drosophil
43	830.6	13.3	3321	13	ADR07314 Full leng
44	829.4	13.2	3064	3	AAF15575 Human pro
45	585.2	9.3	1011	2	AAQ94312 Tyrosine

ALIGNMENTS

RESULT 1

ADI32147
ID ADI32147 standard; CDNA; 6263 BP.

AC ADI32147;
XX

DT 17-JUN-2004 (first entry)
XX

DE Human CDNA #1473.
XX

KW Human; gene; ss; immunological response; immunopathological condition;
KW Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
KW Irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW acute monocytic leukaemia; antiinflammatory; antidiabetic; antitumor;
KW osteopathic; antiarthritic; antirheumatic; cytostatic.
XX Homo sapiens.
XX US6607879-B1.

PD 19-AUG-2003.
XX

PF 09-FEB-1998; 98US-00023655.
XX

PR 09-FEB-1998; 98US-00023655.
XX

PA (INCY-) INCYTE CORP.
XX

PI Cocks BG, Stuart SG, Seilhamer JJ;
XX

DR WPI; 2003-895307/82.
XX

PT A composition comprising a plurality of cDNAs, useful for detecting
PT altered expression of genes in an immunological response or for
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
PT or osteoarthritis.
XX Claim 1; SEQ ID NO 1473; 50pp; English.

PS The invention relates to a composition comprising a plurality of cDNAs
XX for detecting the altered expression of genes in an immunological
CC response. The invention also relates to a method of diagnosing or
CC monitoring the treatment of an immunopathological condition in a sample,

CC comprising obtaining nucleic acids from a sample, contacting the nucleic
CC acids of the sample with an array comprising the plurality of cDNAs under
CC conditions to form one or more hybridisation complexes, detecting the
CC hybridisation complexes and comparing the levels of the detected
CC hybridisation complexes with the level of hybridisation complexes
CC detected in a non-diseased sample, where an altered level of the detected
CC hybridisation complexes correlates with the presence of an
CC immunopathological condition. Also disclosed are an expression profile
CC comprising a microarray and a plurality of detectable complexes and a
CC method for identifying a plurality of polynucleotide probes. The cDNAs
CC are useful as hybridisable array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray can be used in
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
CC ulcerative colitis, hypersensitivity, irritable bowel syndrome,
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
CC identifying agents for the treatment of the diseases. The microarray may
CC also be used in drug discovery and development, toxicological and
CC carcinogenicity studies, forensics or pharmacogenomics. The composition
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
CC genomic fragments. This sequence represents a human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 6263 BP; 1870 A; 1449 C; 1445 G; 1499 T; 0 U; 0 Other;

Query Match 100.0%; Score 6263; DB 11; Length 6263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTAACTCAAGGAGAGCTCTGTGTGAACACCGTGGGATCTAAAGAACAGCTCTGAAAG 60
DB 1 GCTAACTCAAGGAGAGCTCTGTGTGAACACCGTGGGATCTAAAGAACAGCTCTGAAAG 60

QY 61 TGTTCAGCTGAAATTTTCAGATCGGACAGACTCGCTCGGGTCCGGAGGAGCTGATTTCCA 120
DB 61 TGTTCAGCTGAAATTTTCAGATCGGACAGACTCGCTCGGGTCCGGAGGAGCTGATTTCCA 120

QY 121 AGCTGCTCGGACAGCTGCTGCGAAGCTGCGAGATGTTGCGACAGTACGAGGCTGCTGCTG 180
DB 121 AGCTGCTCGGACAGCTGCTGCGAAGCTGCGAGATGTTGCGACAGTACGAGGCTGCTGCTG 180

QY 181 CTGCTCTCTCACTTTCTTCTCCGACAGGATGCTCGACACCTTCCAAAGTTTACACGAACA 240
DB 181 CTGCTCTCTCACTTTCTTCTCCGACAGGATGCTCGACACCTTCCAAAGTTTACACGAACA 240

QY 241 CCCGTTGATCAGACAGGGGTCTCTGGCGAGTTGCTCTTTTCATCTGCCAAGCTACGGGA 300
DB 241 CCCGTTGATCAGACAGGGGTCTCTGGCGAGTTGCTCTTTTCATCTGCCAAGCTACGGGA 300

QY 301 GACCCAGACCTAAATTTGCTGGAAACAAAGGAAAGAAAGTACAGATTTT 360
DB 301 GACCCAGACCTAAATTTGCTGGAAACAAAGGAAAGAAAGTACAGATTTT 360

QY 361 GAGGTATAGAGTTTGAAGTGGTCTGGATCAGTTCTCAGAAATACACCTTACGGACT 420
DB 361 GAGGTATAGAGTTTGAAGTGGTCTGGATCAGTTCTCAGAAATACACCTTACGGACT 420

QY 421 CCGAGGATAGGCGCATTTATGAATGTGTGGCTTCAAAATATGTGGGAGAAATAGTGA 480
DB 421 CCGAGGATAGGCGCATTTATGAATGTGTGGCTTCAAAATATGTGGGAGAAATAGTGA 480

QY 481 TCCACAGACTCAGATTTTGGGAGAGATCAAAATCCAGGGGCTTCCCTACCATTTGAC 540
DB 481 TCCACAGACTCAGATTTTGGGAGAGATCAAAATCCAGGGGCTTCCCTACCATTTGAC 540

QY 541 ATGGGCCCAAGTTGAAGTGGTGTGAGGTACTCGCAGGGCCACCATGCTTTGTCAGCC 600
DB 541 ATGGGCCCAAGTTGAAGTGGTGTGAGGTACTCGCAGGGCCACCATGCTTTGTCAGCC 600

QY 601 AGTGGTAAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC 660
DB 601 AGTGGTAAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC 660

QY 661 AACAAACAATCGTCGTATTATAGCAGTTACGATCAGAAATCTATTGGTGTACACCAATAGA 720
DB 661 AACAAACAATCGTCGTATTATAGCAGTTACGATCAGAAATCTATTGGTGTACACCAATAGA 720

QY 721 GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATAGTGTGTGGCC 780
DB 721 GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATAGTGTGTGGCC 780

QY 781 ACCAAGCAGCGGGGACTCGCTATTCCGCTCTGCGCAATTTATATGTGACAGAGCTGCGA 840
DB 781 ACCAAGCAGCGGGGACTCGCTATTCCGCTCTGCGCAATTTATATGTGACAGAGCTGCGA 840

QY 841 GAAGTTCGCGCTGCCACCAAGATTTCTTATCCCAAGGATTAATCATGAATCATGCGCA 900
DB 841 GAAGTTCGCGCTGCCACCAAGATTTCTTATCCCAAGGATTAATCATGAATCATGCGCA 900

QY 901 GCGGGAAGCGTTAATATACACTGTGTGGCGGTGGGTGACCAATGCTTATGTAAAGTGG 960
DB 901 GCGGGAAGCGTTAATATACACTGTGTGGCGGTGGGTGACCAATGCTTATGTAAAGTGG 960

QY 961 ATGTTGGGGCAGAGATCTGACACCTGAAGATGATATGCAATAGGAGAAATGTGCTA 1020
DB 961 ATGTTGGGGCAGAGATCTGACACCTGAAGATGATATGCAATAGGAGAAATGTGCTA 1020

QY 1021 GAATGATGATGTAAGACAGTCAAGCAATTTACACCTGTGTGCTATGTCAACACTGGT 1080
DB 1021 GAATGATGATGTAAGACAGTCAAGCAATTTACACCTGTGTGCTATGTCAACACTGGT 1080

QY 1081 GTCAATTGAAGCAATAGCAGATCACTGTCAAAGCCTTACCCAAACCTCCAGGAACCTCT 1140
DB 1081 GTCAATTGAAGCAATAGCAGATCACTGTCAAAGCCTTACCCAAACCTCCAGGAACCTCT 1140

QY 1141 GTAGTGACCCAGAGCAGCTACAAGCATCACACTGACGTGGGACTCTGGGAAACCTTGAG 1200
DB 1141 GTAGTGACCCAGAGCAGCTACAAGCATCACACTGACGTGGGACTCTGGGAAACCTTGAG 1200

QY 1201 CTTGTTCTTATTACATAAATTCAGCATAAACCTTAAACCTCTGAGGAACCTTACAAAGAA 1260
DB 1201 CTTGTTCTTATTACATAAATTCAGCATAAACCTTAAACCTCTGAGGAACCTTACAAAGAA 1260

QY 1261 ATTGATGGGTGGCGACCAACGCTACAGTGTGCTGGACTAAGTCTCTACTCGGATTAT 1320
DB 1261 ATTGATGGGTGGCGACCAACGCTACAGTGTGCTGGACTAAGTCTCTACTCGGATTAT 1320

QY 1321 GAATTCAGGGTGTGCTGCTCAATTAACATTGGGGGGGGGCTCCACAGCAACCTGTGCTA 1380
DB 1321 GAATTCAGGGTGTGCTGCTCAATTAACATTGGGGGGGGGCTCCACAGCAACCTGTGCTA 1380

QY 1381 ACACAAACCTCAGAGCAAGCACCATTCCAGTGGCCCGAGGGATGTCCAGGCAACGAATTTG 1440
DB 1381 ACACAAACCTCAGAGCAAGCACCATTCCAGTGGCCCGAGGGATGTCCAGGCAACGAATTTG 1440

QY 1441 AGTTCCAGCACCACTTTTGGTACAGTGGAGGAACTCTGAAGAGCCAAATGGAAGAGTCAA 1500
DB 1441 AGTTCCAGCACCACTTTTGGTACAGTGGAGGAACTCTGAAGAGCCAAATGGAAGAGTCAA 1500

QY 1501 GGATATAGGTTTATATACATGGATCCCACTCAACATGTCAACACTGATGATGATGATG 1560
DB 1501 GGATATAGGTTTATATACATGGATCCCACTCAACATGTCAACACTGATGATGATGATG 1560

QY 1561 AATGTAGTCTCAGACCAAAATCACTTACTATTGGAACCTTGTAGTGGCCCGAGAAACATATTCT 1620
DB 1561 AATGTAGTCTCAGACCAAAATCACTTACTATTGGAACCTTGTAGTGGCCCGAGAAACATATTCT 1620

QY 1621 GTCAAGTCTCTGGCTTTTACCTCAATTTGGAGATGGTCCCTTTTCAAGTGCATACAAAGTC 1680
DB 1621 GTCAAGTCTCTGGCTTTTACCTCAATTTGGAGATGGTCCCTTTTCAAGTGCATACAAAGTC 1680

QY 1681 ATCACTCAGACAGAGTACAGGGCAGCCTAAACCTTCAAGCAGACCTGAGTCTGAA 1740
DB 1681 ATCACTCAGACAGAGTACAGGGCAGCCTAAACCTTCAAGCAGACCTGAGTCTGAA 1740

Qy	1741	ACAAAGTATTTGGCTCTCTGGACACCTCCAGTTCAGATACATTTGCCAACTATGAAGT	1800
Db	1741	ACAAAGTATTTGGCTCTCTGGACACCTCCAGTTCAGATACATTTGCCAACTATGAAGT	1800
Qy	1801	GTCTACAAAGATGGGAGCATGGAGAGCAACGAATTTACATTTGAGCCAGGACATCA	1860
Db	1801	GTCTACAAAGATGGGAGCATGGAGAGCAACGAATTTACATTTGAGCCAGGACATCA	1860
Qy	1861	TATAGGCTGCAAGGACTGAAACCAAGCTTTATATATTTCCGCTGGCTGCAAGCTCC	1920
Db	1861	TATAGGCTGCAAGGACTGAAACCAAGCTTTATATATTTCCGCTGGCTGCAAGCTCC	1920
Qy	1921	CCTCAAGGCTGGGTGCTTCTACTGCAGAAATATCAGCTAGAACCATGCAAGCCG	1980
Db	1921	CCTCAAGGCTGGGTGCTTCTACTGCAGAAATATCAGCTAGAACCATGCAAGCCG	1980
Qy	1981	TCAGCTCCTCCTCAAGACATTTAGTTGACACGACCCCAAGTTCACCTAGTATTTGGTAAGT	2040
Db	1981	TCAGCTCCTCCTCAAGACATTTAGTTGACACGACCCCAAGTTCACCTAGTATTTGGTAAGT	2040
Qy	2041	TGGCAACCTCCACAGTGGAAAAACAGAAATGGCAATTTACCTGAATATCTCCATCAAGTAC	2100
Db	2041	TGGCAACCTCCACAGTGGAAAAACAGAAATGGCAATTTACCTGAATATCTCCATCAAGTAC	2100
Qy	2101	ACTGCAGTGGATGGGAGATGACAAGCCTCAGAGATTTTGGGAATTTCCCTCGGACACT	2160
Db	2101	ACTGCAGTGGATGGGAGATGACAAGCCTCAGAGATTTTGGGAATTTCCCTCGGACACT	2160
Qy	2161	ACCAAAATACCTTTTGGAAACAGCTGGAAAAATGGACTGAATACCGGATCAGCTGACAGCC	2220
Db	2161	ACCAAAATACCTTTTGGAAACAGCTGGAAAAATGGACTGAATACCGGATCAGCTGACAGCC	2220
Qy	2221	CATACAGATGTCGCGCTGGCCCTGAGAGCTTGTCCGTTGTTGATTCGAAACCAATGAAGAT	2280
Db	2221	CATACAGATGTCGCGCTGGCCCTGAGAGCTTGTCCGTTGTTGATTCGAAACCAATGAAGAT	2280
Qy	2281	GTTCCTAGTGTCTCTCGCAAAAGTGCAGGTAGAGCTGTCAACTCAATCTGTTAAA	2340
Db	2281	GTTCCTAGTGTCTCTCGCAAAAGTGCAGGTAGAGCTGTCAACTCAATCTGTTAAA	2340
Qy	2341	GTCTCATGGGCTCACCGCTGCCCAATAAAGATGSCAGATGSCAGATATACAGGTG	2400
Db	2341	GTCTCATGGGCTCACCGCTGCCCAATAAAGATGSCAGATGSCAGATATACAGGTG	2400
Qy	2401	CATTATGTAGGATGGAAAAATGGTAGCCCAAGGCCCAGCCCATGCTGAAAGATGT CATG	2460
Db	2401	CATTATGTAGGATGGAAAAATGGTAGCCCAAGGCCCAGCCCATGCTGAAAGATGT CATG	2460
Qy	2461	CTGGCTGATGCAGTGGGAATTTGATGATATCTACTGAAACATGATCATTTCTGGG	2520
Db	2461	CTGGCTGATGCAGTGGGAATTTGATGATATCTACTGAAACATGATCATTTCTGGG	2520
Qy	2521	CTCCAGCTGAAACTTCTCTACTCCCTCACCGTCAAGCTTACACCAACCAAGGAGATGGT	2580
Db	2521	CTCCAGCTGAAACTTCTCTACTCCCTCACCGTCAAGCTTACACCAACCAAGGAGATGGT	2580
Qy	2581	GCTCGCAGAGCCCAAACTGGTGTCAACCACTGGGCGAGTTCAGGGAAACCTCGGCTT	2640
Db	2581	GCTCGCAGAGCCCAAACTGGTGTCAACCACTGGGCGAGTTCAGGGAAACCTCGGCTT	2640
Qy	2641	GTGATTAACCACTCAGATGAATACTGCTCTTATTCAGTGGCACCCCTCCGGTGGACACA	2700
Db	2641	GTGATTAACCACTCAGATGAATACTGCTCTTATTCAGTGGCACCCCTCCGGTGGACACA	2700
Qy	2701	TTTGGACCTCTTCAGGGTACCGTCTAAAATTTTGGCGCAGGATATGGAGCCACTTACT	2760
Db	2701	TTTGGACCTCTTCAGGGTACCGTCTAAAATTTTGGCGCAGGATATGGAGCCACTTACT	2760
Qy	2761	ACTCTTGAGTTCTCTGAAAAAGAGATCATTCTTACGTACAGATCCCAAGGGAGCA	2820
Db	2761	ACTCTTGAGTTCTCTGAAAAAGAGATCATTCTTACGTACAGATCCCAAGGGAGCA	2820
Qy	2821	TCATACGCTCTCAGGCTCTCAGCAGAAACAAAGTGGGCTTTGGGGAGGATGGTGAAG	2880

Db	2821	TCATACGCTCTCAGGCTCTCAGCAGAAACAAAGTGGGCTTTGGGGAGGATGGTGAAG	2880
Qy	2881	GAGATTTTCCATTTCCAGAAAGTACCAACTGGATTCCCTCAAAAACCTTCACTCAGAAGGC	2940
Db	2881	GAGATTTTCCATTTCCAGAAAGTACCAACTGGATTCCCTCAAAAACCTTCACTCAGAAGGC	2940
Qy	2941	ACCACTTCAACCTCCGTCAGTTATCTTTGGCAACCACTGTCTGGCAGAGAGAAATGGC	3000
Db	2941	ACCACTTCAACCTCCGTCAGTTATCTTTGGCAACCACTGTCTGGCAGAGAGAAATGGC	3000
Qy	3001	ATTATCACCAAGTATACCTTCTTATAGGATATCAACATCCCTCTCTCCCATGGAG	3060
Db	3001	ATTATCACCAAGTATACCTTCTTATAGGATATCAACATCCCTCTCTCCCATGGAG	3060
Qy	3061	CAGCTTATTTGTCAGCTGACACCACTATGACACTCACTGCCTTAAAAACAGATACCACA	3120
Db	3061	CAGCTTATTTGTCAGCTGACACCACTATGACACTCACTGCCTTAAAAACAGATACCACA	3120
Qy	3121	TACGATGTAAAAAGTACGTGCTCATACGAGAAAGGCGCGGGCCATATAGTCCCAAGTGC	3180
Db	3121	TACGATGTAAAAAGTACGTGCTCATACGAGAAAGGCGCGGGCCATATAGTCCCAAGTGC	3180
Qy	3181	CAGTTCAAGGACACTGCTGTGGATCAAGTGTGTTGCAAAAATTTTCATGTCAAAGCAGTA	3240
Db	3181	CAGTTCAAGGACACTGCTGTGGATCAAGTGTGTTGCAAAAATTTTCATGTCAAAGCAGTA	3240
Qy	3241	ATGAAGACTTCCGTTGCTGTCTTGGGAGATTCAGAGAAATATATACTCCGCCATGCT	3300
Db	3241	ATGAAGACTTCCGTTGCTGTCTTGGGAGATTCAGAGAAATATATACTCCGCCATGCT	3300
Qy	3301	TTCAAAATCTTTATGATGATGGAAAAATGGTAGAAGTGGATGGCGGAGCCACACAG	3360
Db	3301	TTCAAAATCTTTATGATGATGGAAAAATGGTAGAAGTGGATGGCGGAGCCACACAG	3360
Qy	3361	AAGTTAATCTCAACCTGAAAGCTGAGAAATCATATTTCTTGTGCTGACAAAATCGTGA	3420
Db	3361	AAGTTAATCTCAACCTGAAAGCTGAGAAATCATATTTCTTGTGCTGACAAAATCGTGA	3420
Qy	3421	AACAGTCTCGTGGCTGACAGCACAGGCTCACGCAAGAGCTGCACAGATGTATTACGT	3480
Db	3421	AACAGTCTCGTGGCTGACAGCACAGGCTCACGCAAGAGCTGCACAGATGTATTACGT	3480
Qy	3481	ACCAAGCTCGCTTCAATGGGAAGACCAACTGGATGGCATGATTCTGTGCAACTGCT	3540
Db	3481	ACCAAGCTCGCTTCAATGGGAAGACCAACTGGATGGCATGATTCTGTGCAACTGCT	3540
Qy	3541	GAAGTACTCTCAATGAGATATAAAGGTTACTACATAATAATTTGTCCTTTCAAGAAA	3600
Db	3541	GAAGTACTCTCAATGAGATATAAAGGTTACTACATAATAATTTGTCCTTTCAAGAAA	3600
Qy	3601	TCTCGCGGAAAATTTTATCAAGCCATGGGAGAGTCCAGATGAAATGGAATAGATGAGCTG	3660
Db	3601	TCTCGCGGAAAATTTTATCAAGCCATGGGAGAGTCCAGATGAAATGGAATAGATGAGCTG	3660
Qy	3661	CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGGAGAGAGTTGAAATTAAG	3720
Db	3661	CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGGAGAGAGTTGAAATTAAG	3720
Qy	3721	CCATATTTGCGCTCACTTTGATGCTTCCCACTGAGTTTCACTCCCTGGGGATGACAAG	3780
Db	3721	CCATATATTCGCTCACTTTGATGCTTCCCACTGAGTTTCACTCCCTGGGGATGACAAG	3780
Qy	3781	CATTATGGTGGATTTTCAAAACAAAGCACTCCAAAGTGGTCAAGAAATATGCTCTTTTGTG	3840
Db	3781	CATTATGGTGGATTTTCAAAACAAAGCACTCCAAAGTGGTCAAGAAATATGCTCTTTTGTG	3840
Qy	3841	TTAGCAGTAAATGGAACATGCAGAGTCTAAGATGTATGAAACCAAGCCCTTACTCGACCCC	3900
Db	3841	TTAGCAGTAAATGGAACATGCAGAGTCTAAGATGTATGAAACCAAGCCCTTACTCGACCCC	3900
Qy	3901	GTGGTGTCAATGATCTGGATCCGAGCCCAATCAGCGATCAAGAAAGGCTTGATCTGG	3960

Db 3061 |||||CAGCTTATGTTCCAGCTGACACCACTATGACACTCACTGGCTTAAACCCAGATACCAACA 3120
Qy 3121 TACGATGTAAAGTACGTGCTCATACGAGCAAAAGGCGCGGCCCATATAGTCCCAAGTGTG 3180
Db 3121 TACGATGTAAAGTACGTGCTCATACGAGCAAAAGGCGCGGCCCATATAGTCCCAAGTGTG 3180
Qy 3181 CAGTTCCAGGACACTGCGCTGTGGATCAAGTGTGGCAAAAATTTTCATGTCAAAAGCAGTA 3240
Db 3181 CAGTTCCAGGACACTGCGCTGTGGATCAAGTGTGGCAAAAATTTTCATGTCAAAAGCAGTA 3240
Qy 3241 ATGAAGACTTCCGCTGTGCTGCTTTGGGAGATTCAGAGAAATTAATACTCCGCGCATGCT 3300
Db 3241 ATGAAGACTTCCGCTGTGCTGCTTTGGGAGATTCAGAGAAATTAATACTCCGCGCATGCT 3300
Qy 3301 TTCAAAAATTTCTTATGATGATGGGAAATGTTAGAGAAATGATGGCGGAGCCACACAG 3360
Db 3301 TTCAAAAATTTCTTATGATGATGGGAAATGTTAGAGAAATGATGGCGGAGCCACACAG 3360
Qy 3361 AAGTTAATGTCAACCTGAAGCTGAGAAATCATATTCATTTGTGCTGACAAAATCGTGGGA 3420
Db 3361 AAGTTAATGTCAACCTGAAGCTGAGAAATCATATTCATTTGTGCTGACAAAATCGTGGGA 3420
Qy 3421 AACAGTGTGCTGGCTGCAGACAGGCTCAGGCAAGAGCTGCACCAAGATGTATTACGT 3480
Db 3421 AACAGTGTGCTGGCTGCAGACAGGCTCAGGCAAGAGCTGCACCAAGATGTATTACGT 3480
Qy 3481 ACCAAGCTGCTCTTCAATTTGGGAAGACCAACTTGGATGGCATGATTACTGTGCAACTGCGCT 3540
Db 3481 ACCAAGCTGCTCTTCAATTTGGGAAGACCAACTTGGATGGCATGATTACTGTGCAACTGCGCT 3540
Qy 3541 GAAGTACTCGCAATAGAGAAATATAAAGGTTACTACATAATAATTTGTGCTTTGAAAGAAA 3600
Db 3541 GAAGTACTCGCAATAGAGAAATATAAAGGTTACTACATAATAATTTGTGCTTTGAAAGAAA 3600
Qy 3601 TCTCGCGGAAATTTATCAAGCCATGGGAGAGTCCAGATGAATGGAATAGATGAGCTG 3660
Db 3601 TCTCGCGGAAATTTATCAAGCCATGGGAGAGTCCAGATGAATGGAATAGATGAGCTG 3660
Qy 3661 CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGGAGAGAGTTGAAATTAAG 3720
Db 3661 CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGGAGAGAGTTGAAATTAAG 3720
Qy 3721 CCATATATGCGCTCACTTTGATGCTCTTCCCACTGAGTTTCACTCGGGGGATGACAAG 3780
Db 3721 CCATATATGCGCTCACTTTGATGCTCTTCCCACTGAGTTTCACTCGGGGGATGACAAG 3780
Qy 3781 CATTTATGTTGATTTACAAACAAGCACTCCAGAGTGTCAAGATATGCTTCTTTG 3840
Db 3781 CATTTATGTTGATTTACAAACAAGCACTCCAGAGTGTCAAGATATGCTTCTTTG 3840
Qy 3841 TTAGCAGTAAATGGAAACATGCAAGATCTAAGATGTATGCAACAGCCCTTACTCCGAGCCCC 3900
Db 3841 TTAGCAGTAAATGGAAACATGCAAGATCTAAGATGTATGCAACAGCCCTTACTCCGAGCCCC 3900
Qy 3901 GTGTGTCAATGTGATCTGGATCCGAGCCCAATCAAGATGAAGAGGCTTGTATCTGG 3960
Db 3901 GTGTGTCAATGTGATCTGGATCCGAGCCCAATCAAGATGAAGAGGCTTGTATCTGG 3960
Qy 3961 GTTGTAGTCTGCTGCTGCTGCTGCTTTTATCATCTGCAATGTCTTCTTCTTCTT 4020
Db 3961 GTTGTAGTCTGCTGCTGCTGCTGCTTTTATCATCTGCAATGTCTTCTTCTTCTT 4020
Qy 4021 TATAAAGGAGAGAGGCGAGCTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAG 4080
Db 4021 TATAAAGGAGAGAGGCGAGCTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAG 4080
Qy 4081 ATCCCTTTCACACCAACCAAGAGCCCTGTGAACTGAGAGGCGCTTAACTTTCAAAACCG 4140
Db 4081 ATCCCTTTCACACCAACCAAGAGCCCTGTGAACTGAGAGGCGCTTAACTTTCAAAACCG 4140
Qy 4141 GGTATGCTAGCCATCTCCATATCCCATCTTGGAACTTGCAGACCACTTGAAGATTG 4200

Db 4141 GGTATGCTAGCCATCTCCATATCCCATCTTGGAACTTGCAGACCACTTGAAGATTG 4200
Qy 4201 AAAGCAAAATGACAACTTTGAAGTTTTTCCAGGAATATGAGTCAATTTGACCTCGGCCAGAG 4260
Db 4201 AAAGCAAAATGACAACTTTGAAGTTTTTCCAGGAATATGAGTCAATTTGACCTCGGCCAGAG 4260
Qy 4261 TTCACTTTGGGAACATTTCAAACTTTGGAAATTAACAAACCAAGAAATAGATACCGGAATGTA 4320
Db 4261 TTCACTTTGGGAACATTTCAAACTTTGGAAATTAACAAACCAAGAAATAGATACCGGAATGTA 4320
Qy 4321 ATCCCATATGATCAATTTCCCGGTTCTCTATCAGCTATAGAGGATCCCAAGGAAGTGAC 4380
Db 4321 ATCCCATATGATCAATTTCCCGGTTCTCTATCAGCTATAGAGGATCCCAAGGAAGTGAC 4380
Qy 4381 TATGTGAATGCCAACTACATAGATGGGTATAGGAAGCAAAATGCTATATTGCAACACAG 4440
Db 4381 TATGTGAATGCCAACTACATAGATGGGTATAGGAAGCAAAATGCTATATTGCAACACAG 4440
Qy 4441 GGATCTCTCCCGGAAACATTTTGGGACTTTTGGAGAAATGATATGGGAACAAAGGATGCC 4500
Db 4441 GGATCTCTCCCGGAAACATTTTGGGACTTTTGGGAGAAATGATATGGGAACAAAGGATGCC 4500
Qy 4501 ACAGTGTCTGATGACAACTAGAAAGAAAGATCAAGGGTGAAGTGTGACCAAGTATTGG 4560
Db 4501 ACAGTGTCTGATGACAACTAGAAAGAAAGATCAAGGGTGAAGTGTGACCAAGTATTGG 4560
Qy 4561 CCTAGCAGAGCAGAGAAACCCACGACTCGTTCAAGTAAACGCTGCTGTGATCTGTGGAG 4620
Db 4561 CCTAGCAGAGCAGAGAAACCCACGACTCGTTCAAGTAAACGCTGCTGTGATCTGTGGAG 4620
Qy 4621 CTGGCCACATATTTGTTTTCGAACATTTTGCACCTTTTACAAAGAAATGGTTCAAGTGAAGAGAGA 4680
Db 4621 CTGGCCACATATTTGTTTTCGAACATTTTGCACCTTTTACAAAGAAATGGTTCAAGTGAAGAGAGA 4680
Qy 4681 GAAGTGAAGCAATTTCCAGTTTCAAGCTGCGCTGCTGATCATGTTTCCAGAACACCCCTACA 4740
Db 4681 GAAGTGAAGCAATTTCCAGTTTCAAGCTGCGCTGCTGATCATGTTTCCAGAACACCCCTACA 4740
Qy 4741 CCTTTCTAGCTTTCTTTACGTAGAGTCAAAACCTGTAAACCTTCCCGATGCTGGTCCGATG 4800
Db 4741 CCTTTCTAGCTTTCTTTACGTAGAGTCAAAACCTGTAAACCTTCCCGATGCTGGTCCGATG 4800
Qy 4801 GTTGTGACATGCGAGTGGGAGTTGGCGGACTGGTTGCTTCATCGTCATAGATGCCATG 4860
Db 4801 GTTGTGACATGCGAGTGGGAGTTGGCGGACTGGTTGCTTCATCGTCATAGATGCCATG 4860
Qy 4861 TTAGAAAGAAATAAGCATGAAACCTGTAGATATTTATGSCCATGTAACTTTAATGAGA 4920
Db 4861 TTAGAAAGAAATAAGCATGAAACCTGTAGATATTTATGSCCATGTAACTTTAATGAGA 4920
Qy 4921 GCCCAGAGGAACCTATATGTTTCAAAACAGAGACCAATACATCTTTTATCCATGATGCACTG 4980
Db 4921 GCCCAGAGGAACCTATATGTTTCAAAACAGAGACCAATACATCTTTTATCCATGATGCACTG 4980
Qy 4981 TTAGAGCAGTGTGCTGTGAAATACCGAAGTGGCAGCTAGAACTTGTATGCTTACATTT 5040
Db 4981 TTAGAGCAGTGTGCTGTGAAATACCGAAGTGGCAGCTAGAACTTGTATGCTTACATTT 5040
Qy 5041 CAGAGCTGACACAAATAGAAACGGGAGAGAAATGTCACAGAAATGGAGCTCGAAATTTAAG 5100
Db 5041 CAGAGCTGACACAAATAGAAACGGGAGAGAAATGTCACAGAAATGGAGCTCGAAATTTAAG 5100
Qy 5101 CGTCTAGCCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCCTTCTTCCATGTAAT 5160
Db 5101 CGTCTAGCCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCCTTCTTCCATGTAAT 5160
Qy 5161 AAATTCAAAAATCGCCTTGTAAATATATGCCATATGAATCCCAAGGGTATGCTCCGAG 5220
Db 5161 AAATTCAAAAATCGCCTTGTAAATATATGCCATATGAATCCCAAGGGTATGCTCCGAG 5220
Qy 5221 CCTATCCGTCAGTGAAGGATCTGATTACATCAATGCCAGTTTATTTGATGGATACAGA 5280
Db 5221 CCTATCCGTCAGTGAAGGATCTGATTACATCAATGCCAGTTTATTTGATGGATACAGA 5280

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2005, 20:04:47 ; Search time 25780 Seconds
(without alignments)
11771.723 Million cell updates/sec

Title: US-10-772-636-63
Perfect score: 6263
Sequence: 1 gctaactcaaggagagcgc.....aagtgaaattttaaaaaagtt 6263

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.on.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6263	100.0	6263	6 AR380928	AR380928 Sequence
2	6263	100.0	6263	9 HUMPTPD	L38929 Homo sapien
3	5657.6	90.3	4584	6 CQ722829	CQ722829 Sequence
4	4584	73.2	5754	9 HSPTPD	X54133 Human HPTP
5	3856	61.6	6562	5 AF197944	AF197944 Xenopus 1
6	2473	39.5	6363	6 AR380334	AR380334 Sequence
7	2473	39.5	6363	9 HSU35234	U35234 Human prote
8	2418.4	38.6	4031	10 AF326559	AF326559 Mus muscu
9	2399.6	38.3	6000	9 HSU40317	U40317 Human prote
10	2398	38.3	4008	10 AF326560	AF326560 Mus muscu
11	2398	38.3	6000	6 AR031690	AR031690 Sequence
12	2398	38.3	6000	6 I61404	I61404 Sequence 6
13	2371	37.9	5723	10 AF300943	AF300943 Mus muscu
14	2365.4	37.8	4555	6 E09890	E09890 Mouse cdna
15	2357.4	37.6	7702	6 CQ871861	CQ871861 Sequence
16	2357.4	37.6	7702	6 AR380791	AR380791 Sequence
17	2357.4	37.6	7702	6 AX658135	AX658135 Sequence
18	2357.4	37.6	7702	9 HSLARR	Y00815 Human mRNA
19	2357.4	37.6	7945	6 CQ496043	CQ496043 Sequence

20	2346.6	37.5	6545	10 RATLARA	L11586 Rat leukocy
21	2321	37.1	5978	10 MMPTPN03	X82288 M.musculus
22	2320.8	37.1	4555	10 MUSMPT03	D13903 Mus musculu
23	2315.8	37.0	6412	9 BC048768	BC048768 Homo sapi
24	2307.8	36.8	6734	6 AX305472	AX305472 Sequence
25	2307.8	36.8	6734	10 MUSPTPT9	D28530 Mouse mRNA
26	2173.8	34.7	6469	10 RATLARPTPA	L11587 Rat leukocy
27	2137.2	34.1	5634	6 CQ721109	CQ721109 Sequence
28	2118.8	33.8	2406	10 BC025145	BC025145 Mus muscu
29	2104.8	33.6	5424	6 CQ722132	CQ722132 Sequence
30	2095.2	32.3	7297	5 AF197945	AF197945 Xenopus 1
31	1926.6	30.8	5435	10 BC057166	BC057166 Mus muscu
32	1555.8	24.8	5953	9 HSM805669	BS537361 Homo sapi
33	1517.6	24.2	4836	6 CQ842367	-CQ842367 Sequence
34	1517.6	24.2	4836	9 AK125325	AK125325 Homo sapi
35	1393.2	22.2	4613	5 CHKCRYP	L32780 Gallus gall
36	1345.8	21.5	4078	6 AR270569	AR270569 Sequence
37	1338.8	21.4	4773	9 HSU41725	U41725 Human prote
38	1338.8	21.4	5490	6 BD183305	BD183305 Novel gen
39	1301.2	20.8	5612	10 BC083188	BC083188 Mus muscu
40	1300	20.8	5690	6 AR060681	AR060681 Sequence
41	1300	20.8	5690	6 AR064169	AR064169 Sequence
42	1298.4	20.7	5383	10 RATLARPTPB	L12329 Rat leukocy
43	1296.8	20.7	4724	10 RATTVRPHOS	L19933 Rattus norv
44	1280.8	20.5	5588	10 BC052462	BC052462 Mus muscu
45	1245.2	19.9	5412	6 BD103687	BD103687 Myocardia

ALIGNMENTS

RESULT 1
AR380928
LOCUS AR380928 6263 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1473 from patent US 6607879.
ACCESSION AR380928
VERSION AR380928.1 GI:40088562
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6263)
Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.
Compositions for the detection of blood cell and immunological
response gene expression
JOURNAL Patent: US 6607879-A 1473 19-AUG-2003;
FEATURES Location/Qualifiers
source 1..6263
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN	Query Match	Score	DB	Length
	100.0%;	6263;	6;	6263;
	Best Local Similarity	100.0%;	Pred. No. 0;	
	Matches 6263;	Conservative	0;	Mismatches
	0;	Indels	0;	Gaps
	0;			
Qy	1	GCTAACTCAAGGAGAGAGCTCTGTGTGAACACCCGCTGGGATCTTAAAGAACCAAGCTCTGAAAG	60	
Db	1	GCTAACTCAAGGAGAGAGCTCTGTGTGAACACCCGCTGGGATCTTAAAGAACCAAGCTCTGAAAG	60	
Qy	61	TGTTCCAGCTGAATTTTCAGATCGGACAGACTCGCTCGGCTCCGGAGGAGTGATTTCCA	120	
Db	61	TGTTCCAGCTGAATTTTCAGATCGGACAGACTCGCTCGGCTCCGGAGGAGTGATTTCCA	120	
Qy	121	AGCTGTCCGCGCAGCTGCTGCCAAGCTGCGAGGATGTCACCTAGCAGCTAGCAGCTCTGCTG	180	
Db	121	AGCTGTCCGCGCAGCTGCTGCCAAGCTGCGAGGATGTCACCTAGCAGCTAGCAGCTCTGCTG	180	
Qy	181	CTGCTCTCTCACTTTTCTTCTCCCGCAGGATGCTGAGACACTTCAAGGTTTACACGAACA	240	
Db	181	CTGCTCTCTCACTTTTCTTCTCCCGCAGGATGCTGAGACACTTCAAGGTTTACACGAACA	240	
Qy	241	CCCGTTGATCAGACAGGGGTCTCTGGCGGAGTTGCTCTTTTCATCTGCCAAGCTACGGGA	300	

Db 241 |||||CCCCGTTGATCAGACAGGGGCTCTGGGGAGTTGCGCTCTTTCATCTGCCAGCTACGGGA 300
Qy 301 GACCAAGACCTAAATTTGTCTGGAAACAAAAGGAAAGAAAGTTCAGCAATTCAGAGATTT 360
Db 301 GACCAAGACCTAAATTTGTCTGGAAACAAAAGGAAAGAAAGTTCAGCAATTCAGAGATTT 360
Qy 361 GAGTAAATAGAGTTTGAAGTCTGGATGGGTCTGGATCAGTCTCAGAAATCAACCCCTTACGGACT 420
Db 361 GAGTAAATAGAGTTTGAAGTCTGGATGGGTCTGGATCAGTCTCAGAAATCAACCCCTTACGGACT 420
Qy 421 CCGAGGATGAGGCAATTTATGAATGTGTGGCTCAAAATATGTGGGAGAAATTAAGTGA 480
Db 421 CCGAGGATGAGGCAATTTATGAATGTGTGGCTCAAAATATGTGGGAGAAATTAAGTGA 480
Qy 481 TCCACCAAGCTCACAGTTTTCGGGAAAGATCAAAATCCCAGGGGCTTCCCTTACCAATGAC 540
Db 481 TCCACCAAGCTCACAGTTTTCGGGAAAGATCAAAATCCCAGGGGCTTCCCTTACCAATGAC 540
Qy 541 ATGGGCCCCAGTTGAAGTGGTTGAGGTACTTCGACGGCCACCATGCTTTGTGACGCC 600
Db 541 ATGGGCCCCAGTTGAAGTGGTTGAGGTACTTCGACGGCCACCATGCTTTGTGACGCC 600
Qy 601 AGTGTAAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC 660
Db 601 AGTGTAAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC 660
Qy 661 AACAAATGTCGTATTAAGCAGTTACGATCAGAAATCTATTTGGTGTACACCAATGA 720
Db 661 AACAAATGTCGTATTAAGCAGTTACGATCAGAAATCTATTTGGTGTACACCAATGA 720
Qy 721 GGAGCCCTTCAGATTGACAGAGTGAAGTCTGACCAAGGAAATATGAGTGTGTGCC 780
Db 721 GGAGCCCTTCAGATTGACAGAGTGAAGTCTGACCAAGGAAATATGAGTGTGTGCC 780
Qy 781 ACCAAAGCGGGGCACTCGCTATTCGCTCTCCGCAATTTATATGTACAGAGAGCTCGGA 840
Db 781 ACCAAAGCGGGGCACTCGCTATTCGCTCTCCGCAATTTATATGTACAGAGAGCTCGGA 840
Qy 841 GAAATTCGCGTGTCCCAACCAAGATTCCTATCCCAACCCATCATGAATCATGCCA 900
Db 841 GAAATTCGCGTGTCCCAACCAAGATTCCTATCCCAACCCATCATGAATCATGCCA 900
Qy 901 GGCGGAGCGTTAATATACCTGTGTGGCGGTGAGGTCACCAATGCCCTTATGTAAGTGG 960
Db 901 GGCGGAGCGTTAATATACCTGTGTGGCGGTGAGGTCACCAATGCCCTTATGTAAGTGG 960
Qy 961 ATGTTGGGGGAGAGATCTGACACCTGAAGATGATGCCAATAGGAAATGTTGCTA 1020
Db 961 ATGTTGGGGGAGAGATCTGACACCTGAAGATGATGCCAATAGGAAATGTTGCTA 1020
Qy 1021 GAACTGAATGATGAAGACAGTACAGCAATTTACCTGTGTGTGTATGTCAACACTGGGT 1080
Db 1021 GAACTGAATGATGAAGACAGTACAGCAATTTACCTGTGTGTGTATGTCAACACTGGGT 1080
Qy 1081 GTCAATGAGCAATAGACAGATCACTGTCAAGCCTTACCACCACTCCAGGAACTCCT 1140
Db 1081 GTCAATGAGCAATAGACAGATCACTGTCAAGCCTTACCACCACTCCAGGAACTCCT 1140
Qy 1141 GTAGTGACCGAGAGCAGCTACAAGCATCACCTGAGGAGCTCTGCGGAAACCTTGAG 1200
Db 1141 GTAGTGACCGAGAGCAGCTACAAGCATCACCTGAGGAGCTCTGCGGAAACCTTGAG 1200
Qy 1201 CCTGTTCTTATTAATTAATTCAGATTAACCTTAAATCTCTGAGGAACTTTACAAGAA 1260
Db 1201 CCTGTTCTTATTAATTAATTCAGATTAACCTTAAATCTCTGAGGAACTTTACAAGAA 1260
Qy 1261 ATTGATGGGTTGGCGACACAGCTTACAGTGTGCTGAGTAAAGTCCCTACTCGGATTA 1320
Db 1261 ATTGATGGGTTGGCGACACAGCTTACAGTGTGCTGAGTAAAGTCCCTACTCGGATTA 1320
Qy 1321 GAAATTCAGGGTTGTGCTCAATAACATTTGGGCGGGGCTTCCAGCGAACTGTGCTA 1380

Db 1321 GAAATTCAGGGTTGTGCTCAATAAATTCATTGGGGGGGCTTCCAGCGAACTGTGCTA 1380
Qy 1381 ACACAAACCTCAGAGCAGACCATTCAGTGTCCCGAGGGATGTCCAGGACCGAATGTTG 1440
Db 1381 ACACAAACCTCAGAGCAGACCATTCAGTGTCCCGAGGGATGTCCAGGACCGAATGTTG 1440
Qy 1441 AGTTTCAGACCACTTTTGGTACAGTGAAGGAACCTTGAAGAGCCAAATGACACATCCAA 1500
Db 1441 AGTTTCAGACCACTTTTGGTACAGTGAAGGAACCTTGAAGAGCCAAATGACACATCCAA 1500
Qy 1501 GGTATAGAGTTTATATACAATGGATCCCACTCAACATGTCAACAACTGGATGAACAC 1560
Db 1501 GGTATAGAGTTTATATACAATGGATCCCACTCAACATGTCAACAACTGGATGAACAC 1560
Qy 1561 AATGTAGCTGACAGCCAAATCACTATTTGGCAACTTAGTGCCCGCAGAAACATATTTCT 1620
Db 1561 AATGTAGCTGACAGCCAAATCACTATTTGGCAACTTAGTGCCCGCAGAAACATATTTCT 1620
Qy 1621 GTCAAAAGTCTGCTTTTACCTCAATTTGGAGATGGTCCCTTTCAAGTGACATACAAGTC 1680
Db 1621 GTCAAAAGTCTGCTTTTACCTCAATTTGGAGATGGTCCCTTTCAAGTGACATACAAGTC 1680
Qy 1681 ATCACTCAGACAGGAGTACAGGGCAGCCACTAAACTTCAAAAGCAGAACTGAGTCTGAA 1740
Db 1681 ATCACTCAGACAGGAGTACAGGGCAGCCACTAAACTTCAAAAGCAGAACTGAGTCTGAA 1740
Qy 1741 ACAAGTATTTTGTCTCTTGGACACCTCCAGCTTCAGATACCATTTGCCAATATGAAGTC 1800
Db 1741 ACAAGTATTTTGTCTCTTGGACACCTCCAGCTTCAGATACCATTTGCCAATATGAAGTC 1800
Qy 1801 GTCTCAAAAGATGGGAGCATGGAGAGCAACGAATTAACCATTTGAGCCAGGACATCA 1860
Db 1801 GTCTCAAAAGATGGGAGCATGGAGAGCAACGAATTAACCATTTGAGCCAGGACATCA 1860
Qy 1861 TATAGGCTGCAAGGACTGAAAACCAAAACAGCTTATCTATTTCCGTCTGGCTGCAAGCTCC 1920
Db 1861 TATAGGCTGCAAGGACTGAAAACCAAAACAGCTTATCTATTTCCGTCTGGCTGCAAGCTCC 1920
Qy 1921 CCTCAAGGCTGGGTCTTCTACTGCAAGAAATATCAGTAGAACCATTCAGTCAAGGCCG 1980
Db 1921 CCTCAAGGCTGGGTCTTCTACTGCAAGAAATATCAGTAGAACCATTCAGTCAAGGCCG 1980
Qy 1981 TCAGCTCTCTCTCAAGACATTTAGTTCACACAGCCCAAGTTCCACTAGTATTTTGGTAAGT 2040
Db 1981 TCAGCTCTCTCTCAAGACATTTAGTTCACACAGCCCAAGTTCCACTAGTATTTTGGTAAGT 2040
Qy 2041 TGGCAACCTCCACAGTGGGAAAAACAGAAATGGCAATTAATCACTGAATCTCCATCAAGTAC 2100
Db 2041 TGGCAACCTCCACAGTGGGAAAAACAGAAATGGCAATTAATCACTGAATCTCCATCAAGTAC 2100
Qy 2101 ACTGCAGTGGATGGGAGATGACAGCCTCAGAGATTTTGGGAATTTCTTCGGACACT 2160
Db 2101 ACTGCAGTGGATGGGAGATGACAGCCTCAGAGATTTTGGGAATTTCTTCGGACACT 2160
Qy 2161 ACCAAATACCTTTTGGAAACAGCTGGAAAAATGGAATACCGGATCACTGTGACAGCC 2220
Db 2161 ACCAAATACCTTTTGGAAACAGCTGGAAAAATGGAATACCGGATCACTGTGACAGCC 2220
Qy 2221 CATACAGATGTCGGCTGCGCCCTGAGAGCTGTGCGGTGTTGATTCGAAACCAATGAAGAT 2280
Db 2221 CATACAGATGTCGGCTGCGCCCTGAGAGCTGTGCGGTGTTGATTCGAAACCAATGAAGAT 2280
Qy 2281 GTTCTTAGTGTCTCTCTCCGCAAGTGCAGGTAGAGCTGTCAACTCAACTCTGTTAAA 2340
Db 2281 GTTCTTAGTGTCTCTCTCCGCAAGTGCAGGTAGAGCTGTCAACTCAACTCTGTTAAA 2340
Qy 2341 GTCTCATGGGCTCACCCGTGCCAATAAAGACAGATGGCCAGATAAAGAGGATACAGGTG 2400
Db 2341 GTCTCATGGGCTCACCCGTGCCAATAAAGACAGATGGCCAGATAAAGAGGATACAGGTG 2400
Qy 2401 CATTTATGAGGATGGAATAATGTTGAGCCCAAGGGCCAGCCCATGCTGGAAGATGTCATG 2460
Db 2401 CATTTATGAGGATGGAATAATGTTGAGCCCAAGGGCCAGCCCATGCTGGAAGATGTCATG 2460

QY 2461 CTGGCTGATGCACAGTGGGAATTTGATGATCTACTGAAACATGATCATTTCTGGG 2520
DB 2461 CTGGCTGATGCACAGTGGGAATTTGATGATCTACTGAAACATGATCATTTCTGGG 2520
QY 2521 CTCACGCTGAAACTTCTCTACTCCCTCACCGTCAAGCTTCAAGGAGATGATGAT 2580
DB 2521 CTCACGCTGAAACTTCTCTACTCCCTCACCGTCAAGCTTCAAGGAGATGATGAT 2580
QY 2581 GCTCGACAGCAGCCCAACTGCTGCTCAACCTGCGGAGTTCACAGGAAACCTCGGCTT 2640
DB 2581 GCTCGACAGCAGCCCAACTGCTGCTCAACCTGCGGAGTTCACAGGAAACCTCGGCTT 2640
QY 2641 GTGATTAACCACTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
DB 2641 GTGATTAACCACTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
QY 2701 TTTGGACCTTTCAGGCTTACCGTCTAAATTTGGCCGCAAGGATGATGAGCCTTACT 2760
DB 2701 TTTGGACCTTTCAGGCTTACCGTCTAAATTTGGCCGCAAGGATGATGAGCCTTACT 2760
QY 2761 ACTCTGAGTCTCTGAAAGAGAGATCACTTTACAGCTACAGATCCCAAGGGAGCA 2820
DB 2761 ACTCTGAGTCTCTGAAAGAGAGATCACTTTACAGCTACAGATCCCAAGGGAGCA 2820
QY 2821 TCATACGCTTTCAGGCTTTCAGGCTTTCAGGCTTTCAGGCTTTCAGGCTTTCAGG 2880
DB 2821 TCATACGCTTTCAGGCTTTCAGGCTTTCAGGCTTTCAGGCTTTCAGGCTTTCAGG 2880
QY 2881 GAGATTTCCATTCAGAGAGAGTCACTGATGATGATGATGATGATGATGATGATGAT 2940
DB 2881 GAGATTTCCATTCAGAGAGAGTCACTGATGATGATGATGATGATGATGATGATGAT 2940
QY 2941 ACCACTTCAACCTCGTCCAGTATCTTTGGCAACCACTGCTGCTGGCAGAGAGAAATGGC 3000
DB 2941 ACCACTTCAACCTCGTCCAGTATCTTTGGCAACCACTGCTGCTGGCAGAGAGAAATGGC 3000
QY 3001 ATTATCAGAGTATACCTCTTTTATAGGATATCAACATCCCTCTCCGATGGAG 3060
DB 3001 ATTATCAGAGTATACCTCTTTTATAGGATATCAACATCCCTCTCCGATGGAG 3060
QY 3061 CAGCTTATGTTTCAGCTGACACCACTATGATGATGATGATGATGATGATGATGATGAT 3120
DB 3061 CAGCTTATGTTTCAGCTGACACCACTATGATGATGATGATGATGATGATGATGATGAT 3120
QY 3121 TACGATGTAAGTACGCTCATAAGCAAGGCGCCCGGCGCCATATAGTCCAGTGTC 3180
DB 3121 TACGATGTAAGTACGCTCATAAGCAAGGCGCCCGGCGCCATATAGTCCAGTGTC 3180
QY 3181 CAGTTTACGAGCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
DB 3181 CAGTTTACGAGCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
QY 3241 ATGAAGACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
DB 3241 ATGAAGACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
QY 3301 TTCAAAATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
DB 3301 TTCAAAATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
QY 3361 AAGTTAATTTGCAACCTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACT 3420
DB 3361 AAGTTAATTTGCAACCTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACT 3420
QY 3421 AACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
DB 3421 AACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
QY 3481 ACCAAGCTGCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
DB 3481 ACCAAGCTGCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540

QY 3541 GAAGTACCTGCAATGAGAAATATAAAGTTTACTACATAATAATTTGTCCTTTGAGAAA 3600
DB 3541 GAAGTACCTGCAATGAGAAATATAAAGTTTACTACATAATAATTTGTCCTTTGAGAAA 3600
QY 3601 TCTCGGGAAATTTTATCAAGCCATGGAGAGTCCAGATGAAATGGAATTTAGATGAGCTG 3660
DB 3601 TCTCGGGAAATTTTATCAAGCCATGGAGAGTCCAGATGAAATGGAATTTAGATGAGCTG 3660
QY 3661 CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGGAGAGAAATTTGAAATTAAG 3720
DB 3661 CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGGAGAGAAATTTGAAATTAAG 3720
QY 3721 CCATATATGCGCTCACTTTGATGTCCTTCCCACTGAGTTTCACTCCGCGGAGTACAAAG 3780
DB 3721 CCATATATGCGCTCACTTTGATGTCCTTCCCACTGAGTTTCACTCCGCGGAGTACAAAG 3780
QY 3781 CATTATGCTGATTTTACAAACAGCAACTCCAAAGTGGTCAAGAAATATGCTTCTTTG 3840
DB 3781 CATTATGCTGATTTTACAAACAGCAACTCCAAAGTGGTCAAGAAATATGCTTCTTTG 3840
QY 3841 TTAGCAGTAATGGAACATGAGAGTCTAAGATGATGCAACAGCCCTTATCTCGACCCC 3900
DB 3841 TTAGCAGTAATGGAACATGAGAGTCTAAGATGATGCAACAGCCCTTATCTCGACCCC 3900
QY 3901 GTGCTGTAATGGAATCTGGATCCGACCCATCAGGATGAAAGAGGCTTGTCTGG 3960
DB 3901 GTGCTGTAATGGAATCTGGATCCGACCCATCAGGATGAAAGAGGCTTGTCTGG 3960
QY 3961 GTTGTAGCTCTGCTCTGAGTGGTCTTTATCATCTGCTGCTTATCTTCTTCTT 4020
DB 3961 GTTGTAGCTCTGCTCTGAGTGGTCTTTATCATCTGCTGCTTATCTTCTTCTT 4020
QY 4021 TATAAAGGAGAGGCGAGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGAG 4080
DB 4021 TATAAAGGAGAGGCGAGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGAG 4080
QY 4081 ATCCCTTTCACACCAACCAAGACCTGTAAGCTGAGGCGCTTAACTTTCAACACCG 4140
DB 4081 ATCCCTTTCACACCAACCAAGACCTGTAAGCTGAGGCGCTTAACTTTCAACACCG 4140
QY 4141 GGTATGCTAGCCATCTCCAAATACCCATCTTGAACTTTGAGAGCAGCATGAAAGATG 4200
DB 4141 GGTATGCTAGCCATCTCCAAATACCCATCTTGAACTTTGAGAGCAGCATGAAAGATG 4200
QY 4201 AAAGCAATGACAACTTGAAGTTTCCAGGAATATGATGCAATTTGACCTTGGCCAGCAG 4260
DB 4201 AAAGCAATGACAACTTGAAGTTTCCAGGAATATGATGCAATTTGACCTTGGCCAGCAG 4260
QY 4261 TTCACTTTGGGAACTTCAAACTTGGAAATTAACAAACCAAGATAGATGCGGAATGTA 4320
DB 4261 TTCACTTTGGGAACTTCAAACTTGGAAATTAACAAACCAAGATAGATGCGGAATGTA 4320
QY 4321 ATGCTATGATCATTTCCCGGTTCTCTATCAGCTATAGAGGATCCCAAGGAGTAC 4380
DB 4321 ATGCTATGATCATTTCCCGGTTCTCTATCAGCTATAGAGGATCCCAAGGAGTAC 4380
QY 4381 TATGTGAATCCCACTACATAGATGGTATAGGAAGCAAAATGCTTATTTGCAACACAG 4440
DB 4381 TATGTGAATCCCACTACATAGATGGTATAGGAAGCAAAATGCTTATTTGCAACACAG 4440
QY 4441 GGATCTCTCCCGGAACTTTGGGAGCTTTTGGAGATGATGAGGAAACCAAGGAGTGC 4500
DB 4441 GGATCTCTCCCGGAACTTTGGGAGCTTTTGGAGATGATGAGGAAACCAAGGAGTGC 4500
QY 4501 ACAGTTCTCATGATGACAAACTAGAGAAAGATCAAGGCTGAGTGTGACAGTATGG 4560
DB 4501 ACAGTTCTCATGATGACAAACTAGAGAAAGATCAAGGCTGAGTGTGACAGTATGG 4560
QY 4561 CCTAGCAGAGGCAAGAAACCCAGGCTGCTTCAAGTAAAGCTGCTTGTGAGTGTGAG 4620
DB 4561 CCTAGCAGAGGCAAGAAACCCAGGCTGCTTCAAGTAAAGCTGCTTGTGAGTGTGAG 4620
QY 4621 CTGGCCACATATTTGTTGCTTCAACATTTGCACTTTTACAGAAATGGTTTCAAGGAGAG 4680

```
Db 4621 |||||CTGCCACATATTGTGTTGCGNACATTTGCACTTTTACAGAAATGGTTCAAGTGAGAGAGA 4680
Qy 4681 GAAGTGAACAATTCAGTTTCAACCGCTGCGCTGATCATGGTGTTCAGAAACCCCTACA 4740
Db 4681 GAAGTGAACAATTCAGTTTCAACCGCTGCGCTGATCATGGTGTTCAGAAACCCCTACA 4740
Qy 4741 CCTTTCTAGCTTTCTTACGTFAGAGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATG 4800
Db 4741 CCTTTCTAGCTTTCTTACGTFAGAGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATG 4800
Qy 4801 GTTGTGCACTGCAGTGGGAGTTGGCGGAGTGGTTGCTTTCATGCTCATAGATGCCATG 4860
Db 4801 GTTGTGCACTGCAGTGGGAGTTGGCGGAGTGGTTGCTTTCATGCTCATAGATGCCATG 4860
Qy 4861 TTAGAAAGAAATAAGCATGAAACCTGTGATATTTATGGCCATGTAATTAAGAGA 4920
Db 4861 TTAGAAAGAAATAAGCATGAAACCTGTGATATTTATGGCCATGTAATTAAGAGA 4920
Qy 4921 GCCCAGAGGAACCTATATGTTTCAACACAGAGACCAATACATCTTTATCCATGATGCACTG 4980
Db 4921 GCCCAGAGGAACCTATATGTTTCAACACAGAGACCAATACATCTTTATCCATGATGCACTG 4980
Qy 4981 TTAGAACGAGTGACTTGTGGAATATACCGAAGTGCAGCTAGAAAATTGTATGCTCATATT 5040
Db 4981 TTAGAACGAGTGACTTGTGGAATATACCGAAGTGCAGCTAGAAAATTGTATGCTCATATT 5040
Qy 5041 CAGAGCTGACACAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
Db 5041 CAGAGCTGACACAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
Qy 5101 CGTCTAGCCAGCTCAGAGCTCAGACCTCAGAGTTTATCAGTGCCATCTTCCATGTAAT 5160
Db 5101 CGTCTAGCCAGCTCAGAGCTCAGACCTCAGAGTTTATCAGTGCCATCTTCCATGTAAT 5160
Qy 5161 AAATTCAAAATCGCCTTGTAAATATTTATGCAATATGCAATATGCAATATGCAATATGCAAT 5220
Db 5161 AAATTCAAAATCGCCTTGTAAATATTTATGCAATATGCAATATGCAATATGCAATATGCAAT 5220
Qy 5221 CCTATCGTGAGTGAAGAGATCTGATTTACATGCGCAGTTTATTTATGATGATACAGA 5280
Db 5221 CCTATCGTGAGTGAAGAGATCTGATTTACATGCGCAGTTTATTTATGATGATACAGA 5280
Qy 5281 CAACAGAAAGCTACATCGCTTACCCAGGGGCTTGGCAGAGACCTGGAAGACTTCTGG 5340
Db 5281 CAACAGAAAGCTACATCGCTTACCCAGGGGCTTGGCAGAGACCTGGAAGACTTCTGG 5340
Qy 5341 CGGATGCTCTGGGAACACAATTCACCATATGTTGATGCTCACCAAGCTCGGTGAAATG 5400
Db 5341 CGGATGCTCTGGGAACACAATTCACCATATGTTGATGCTCACCAAGCTCGGTGAAATG 5400
Qy 5401 GGCAGAGAAATGTCACCAATATCGCCAGCAGAAACGGTGTGCAAGATACCATGATCTTT 5460
Db 5401 GGCAGAGAAATGTCACCAATATCGCCAGCAGAAACGGTGTGCAAGATACCATGATCTTT 5460
Qy 5461 GTTGTAGATCCATGCTGAGTACAAATGCGCAGATATATCTTCAAGGAAATCAAGGTC 5520
Db 5461 GTTGTAGATCCATGCTGAGTACAAATGCGCAGATATATCTTCAAGGAAATCAAGGTC 5520
Qy 5521 ACAGATGCCAGGACGGCCAGTCCCGAACAGTAAAGGAGTTTCCAGTTTCACTGACTGCCA 5580
Db 5521 ACAGATGCCAGGACGGCCAGTCCCGAACAGTAAAGGAGTTTCCAGTTTCACTGACTGCCA 5580
Qy 5581 GAGCAAGAGTGCACAAAGTCCGGAGAGGATTTATTCATCTCATCGCCAGTCCATAAA 5640
Db 5581 GAGCAAGAGTGCACAAAGTCCGGAGAGGATTTATTCATCTCATCGCCAGTCCATAAA 5640
Qy 5641 ACAAAGAAAGTTTGGCAGAGTGGACCCATTTTCACTTCAAGTCCAGCGGGGCTTGGGA 5700
Db 5641 ACAAAGAAAGTTTGGCAGAGTGGACCCATTTTCACTTCAAGTCCAGCGGGGCTTGGGA 5700
Qy 5701 AGAATCGAGTCTTCAATACCGCTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGAGTT 5760
Db 5701 AGAATCGAGTCTTCAATACCGCTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGAGTT 5760
```

```
Db 5701 AGAATCGAGTCTTCAATACCGCTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGAGTT 5760
Qy 5761 GTAGATATCTTCCAGACTGTCAAAATGTTAAGAACACACAGCAGCAGCTATGGTACAGCA 5820
Db 5761 GTAGATATCTTCCAGACTGTCAAAATGTTAAGAACACACAGCAGCAGCTATGGTACAGCA 5820
Qy 5821 GAGGATCAATATCAGTTTCTTATCGTGGCGCACTAGAGTACCTGGGCGAGCTTTTGACCAC 5880
Db 5821 GAGGATCAATATCAGTTTCTTATCGTGGCGCACTAGAGTACCTGGGCGAGCTTTTGACCAC 5880
Qy 5881 TATGCAACGTAGAAAACCCCTGACCCATCTTGGATTTTACTACAGGCGCTTCAATATCCA 5940
Db 5881 TATGCAACGTAGAAAACCCCTGACCCATCTTGGATTTTACTACAGGCGCTTCAATATCCA 5940
Qy 5941 TGGAGTCTCTTCTGAGCCATACAGGCGACTTGAAGAGTCTCTTAACTTCTAGCTAACA 6000
Db 5941 TGGAGTCTCTTCTGAGCCATACAGGCGACTTGAAGAGTCTCTTAACTTCTAGCTAACA 6000
Qy 6001 ACTACTTAGTGGGACTATTACACACAAAACAAATTAATAACAAATTAATCCAGGTGGACC 6060
Db 6001 ACTACTTAGTGGGACTATTACACACAAAACAAATTAATAACAAATTAATCCAGGTGGACC 6060
Qy 6061 AGAATTTCTTTCATGCGCCCTTCCACCATCTGCTCATATAACATTTTAGGGGCCA 6120
Db 6061 AGAATTTCTTTCATGCGCCCTTCCACCATCTGCTCATATAACATTTTAGGGGCCA 6120
Qy 6121 AGGGAGGGAGTCTTTTAAAGAAAGTCTTGAATTTAGTTAGTTTAAAGATAC 6180
Db 6121 AGGGAGGGAGTCTTTTAAAGAAAGTCTTGAATTTAGTTAGTTTAAAGATAC 6180
Qy 6181 TGCTGACCTGCTCTTCAATTTCTAACTGTGTAAACTTTTTTAAACAAATGTATCATTCG 6240
Db 6181 TGCTGACCTGCTCTTCAATTTCTAACTGTGTAAACTTTTTTAAACAAATGTATCATTCG 6240
Qy 6241 ATAAAGTGAATTTTAAAGAGTT 6263
Db 6241 ATAAAGTGAATTTTAAAGAGTT 6263
```

RESULT 2

```
HUMPTPD HUMPTPD 6263 bp mRNA linear PRI 09-APR-1995
LOCUS Homo sapiens protein tyrosine phosphatase delta mRNA, complete cds.
DEFINITION L38929
ACCESSION L38929.1 GI:755652
VERSION protein tyrosine phosphatase delta; transmembrane protein.
KEYWORDS Homo sapiens
SOURCE Homo sapiens
ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 6263)
AUTHORS Pulido,R., Krueger,N.X., Serra-Pages,C., Saito,H. and Streuli,M.
TITLE Molecular characterization of the human transmembrane
protein-tyrosine phosphatase delta. Evidence for tissue-specific
expression of alternative human transmembrane protein-tyrosine
phosphatase delta isoforms
J. Biol. Chem. 270 (12), 6722-6728 (1995)
JOURNAL 95204468
MEDLINE 7896816
COMMENT Original source text: Homo sapiens cdna to mRNA.
FEATURES Location/Qualifiers
source 1. .6263
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .6263
/codon_start=1
/product="protein tyrosine phosphatase delta"
/protein_id="AAC41749.1"
/db_xref="GI:755653"
/translation="MVHAKLLLLLLTFFLRDTAETPPRFRTRTPVDOTGSGGVASF1
CQATGPRPKIVWNKKGVNQRFVIEFDDGGSVLRIQLRTPRDEAIYECVSN
```

NVGEISVSTRLTVLREDQI PRGFPTIDMGPQLKVVERTRTATMLCAASGNPDPEITWF
 KDPLFVDSNNNGRIQSRSESGTPIRGALQIEQSBESDQKTECVATNSAGTRY5
 APANLYVRELREVRVPRFSIPTPHIMPSSNITCVAGSPMSPPVKNMGLABDL
 TPEDDMPTGRNVLNLEVRQANYSITCVAMSTLGVIEAIAQITVKALPPPGTPVVTES
 TATSITLWDSGNPEPVSYYIIQHKPNSEELYKEIDGVATTRYSVAGLSYPSYDEFR
 VVAVNIGRGPBSPVLQTSQAPSSAPRDVQARMLSSITLIVQWKEPBPNGIOIQ
 YRVVYTMPTQHOVNNMKNHVADSOITTIIGNLPQKTYSVKVLAFSTLGDPLSDIQ
 VITQGVGPQLNFKABESETSLSSWTTPSRSDTIANYELVYKDGHEGSEQRITIEP
 STLSYQLQKSNLSYIFRLAARSPOGLGASTAEISARTMQSPSPAPODISCTSPST
 SILVSWPPPEKQNGIITEYSIKYTAVDGEDDPHEILGIPSDTKYLLLEQLEKWE
 YRIITATHTDVGPESESILVIRNEDVPSGPRKVEAVEAVNSTSVKWSRSPVKNQ
 HQIRGYOHVYRMENGSPKQPMKLDVLAOWEPDDTTEHDMISGLQBPETSYSL
 TVTA VTTKGDGARSKPLVSTDTGAPGKPRLVINHOMNTALIOHPHPVDPTFGLQY
 BLKSRGDFEPLTLEPSEKEDHFTATDIHKGASVYFRLSARNKVGFGEEMKEISIP
 BEVPTQFQNLHSEGTSSVOLSWQPPDPIHKAASVYFRLSARNKVGFGEEMKEISIP
 PADTMTLTKDPTDYDVKVRHASTSGPGPYSPSQVQFRLPVDVQFKNPVKAVM
 KTVLLSWWEIIPENTNSAMPFKLIDYDGHQVEVDGRATQKLIIVNLKPKSYSFVLNVR
 GNSAGLQKRWTPADVLRTKPAFIGKTNLDGMITVOLPEVPANENIKGYIIIVP
 LKSKRGKFIKPWESPDEMLDELLEKESIKRERSIRYGREVELKPYIAHPDVLTEFT
 LGDDKHGFTNKQLQSQEYVFFVLAWMEHAKMYATSPYSDPVMSMDLPQPTD
 BEEGLIWVGPVLAWVFIICILVIALLLYKRRASDSRKSIPNNKEIPSHHPTDVE
 LRRLNFQTPGMAHPPIPIBLADHIERLKANDNLKFSQYESIDPGQQTWESNLE
 VNKPNRYAIADHSLVLLSAIEGIPGSDYVNANYIDGYRKQNAVITQGSLEPTF
 GDFWRMIWEORSATVMMTKLEERSRVKCDYWPSSRGTEHGLVQVTLITDVELATVC
 VRTALYKNGSSEKREVRQFOTAPDQGVPEHPTPLAFLRRVKTCTNPPDAGPMVH
 CSAGVGTGCTIVIDAMLERIKHEKTVDIYGHVTLMRQRNMVQTEQYIFIDHALL
 EAVTCGNTVEPARNLAIYIKQLTOETGENTVMELEPKRLASKAHSTRFISANLPC
 NFKNRLNIMPYESTYRCLQIRGVEGSDYINAFIDGYRQKAYTATQOPLAEYTE
 DFMRLMHNHSTIIVMMITKLREMGREKCHOYPAERSARYQYFVVDPMAYNMPQYTL
 REFKVTARDQSKRTVQFOTDWPQGVKPSGEGFIDFIGOVHKTKEQFGQDGPISV
 HCSAGVGTGCTVIFITLSIVLERMYEGVVDIFQTVKMLRTQRPAMVQTEQYQFSYRA
 LEYLGSFPHYAT"

6263

polYA_site
 ORIGIN

Query Match 100.0%; Score 6263; DB 9; Length 6263;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 6263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GCTAACTCAAGGAGAGCGTCTGGTGAACACCCGCTGGGATCTAAAGAACCAAGCTCTGAAAG	60
Db	1	GCTAACTCAAGGAGAGCGTCTGGTGAACACCCGCTGGGATCTAAAGAACCAAGCTCTGAAAG	60
Qy	61	TGTTCCAGCTCAAAATTTTCAGATCGGACAGACTCGTGGCGCTCGGAGGAGATGATCCA	120
Db	61	TGTTCCAGCTCAAAATTTTCAGATCGGACAGACTCGTGGCGCTCGGAGGAGATGATCCA	120
Qy	121	AGCTGCTCGGACAGCTGCTGCCAAGCTGACAGGATGGTGACAGTGGTCCAGGCTCTGCTG	180
Db	121	AGCTGCTCGGACAGCTGCTGCCAAGCTGACAGGATGGTGACAGTGGTCCAGGCTCTGCTG	180
Qy	181	CTGCTCTCTCACTTTCTTCTCCGACCGGATGCTGAGACACCTCCAAGGTTTACACGAACA	240
Db	181	CTGCTCTCTCACTTTCTTCTCCGACCGGATGCTGAGACACCTCCAAGGTTTACACGAACA	240
Qy	241	CCCGTTGATCAGACAGGGGCTCTGCGGAGTGGCTCTTTCATCTCGCAAGCTACCGGA	300
Db	241	CCCGTTGATCAGACAGGGGCTCTGCGGAGTGGCTCTTTCATCTCGCAAGCTACCGGA	300
Qy	301	GACCAAGACCTAAAATTTGCTCGGAAACAAAAGGAAAGAAAGTCAAGATTCAGATTT	360
Db	301	GACCAAGACCTAAAATTTGCTCGGAAACAAAAGGAAAGAAAGTCAAGATTCAGATTT	360
Qy	361	GAGTAATAGAGTTTGAAGTGGGCTCTGGATCAGTTCTCAGAAATACACCTTACGGACT	420
Db	361	GAGTAATAGAGTTTGAAGTGGGCTCTGGATCAGTTCTCAGAAATACACCTTACGGACT	420
Qy	421	CCGAGGATGAGGCCATTTATGAATGTGGCTCTAAATTAATGTGGAGAAATAAGTGTA	480
Db	421	CCGAGGATGAGGCCATTTATGAATGTGGCTCTAAATTAATGTGGAGAAATAAGTGTA	480
Qy	481	TCCACAGACTCAGATTTTGGGGAAGATCAAAATTCACAGGGGCTTCCCTACCATTTGAC	540
Db	481	TCCACAGACTCAGATTTTGGGGAAGATCAAAATTCACAGGGGCTTCCCTACCATTTGAC	540

Db	481	TCCACAGACTCAGATTTTGGGGAAGATCAAAATTCACAGGGGCTTCCCTACCATTTGAC	540
Qy	541	ATGGGCCCCACAGTTTGAAGGTGGTTGAGCGTACTCGCACGGCCACCATGCTTTGTGCAGCC	600
Db	541	ATGGGCCCCACAGTTTGAAGGTGGTTGAGCGTACTCGCACGGCCACCATGCTTTGTGCAGCC	600
Qy	601	AGTGGTAATCCGGATCCGAAATCAGTTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC	660
Db	601	AGTGGTAATCCGGATCCGAAATCAGTTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC	660
Qy	661	AACAACAAATGTCGTATTAAGCAGTTACGATCAGATCTATTTGGTGGTACACCAATAAGA	720
Db	661	AACAACAAATGTCGTATTAAGCAGTTACGATCAGATCTATTTGGTGGTACACCAATAAGA	720
Qy	721	GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGAGTGTGTGCC	780
Db	721	GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGAGTGTGTGCC	780
Qy	781	ACCAACAGCGGGGCACTCGCTATTTCGCTCTCGCTCCCAATTTATATGTCAGAGAGCTGCGA	840
Db	781	ACCAACAGCGGGGCACTCGCTATTTCGCTCTCGCTCCCAATTTATATGTCAGAGAGCTGCGA	840
Qy	841	GAAGTTCGCGGTGCCCAACCAAGATTCTCTATCCCAACCCACTAATCATGAATCATGCCA	900
Db	841	GAAGTTCGCGGTGCCCAACCAAGATTCTCTATCCCAACCCACTAATCATGAATCATGCCA	900
Qy	901	GGCGGAAGCGTTAATATCACTGTGTGGCGGTGGGGTCAACCAATGCTTATGTAAGTGG	960
Db	901	GGCGGAAGCGTTAATATCACTGTGTGGCGGTGGGGTCAACCAATGCTTATGTAAGTGG	960
Qy	961	ATGTTGGGGGAGAGATCTGACACCTGGAAGATGATATGCCAATAGGAAGAAATGTGCTA	1020
Db	961	ATGTTGGGGGAGAGATCTGACACCTGGAAGATGATATGCCAATAGGAAGAAATGTGCTA	1020
Qy	1021	GAATGTAATGATGTAAGACAGTCAAGCAATTTACACCTGTGTGTGTATGTCAACACTGGT	1080
Db	1021	GAATGTAATGATGTAAGACAGTCAAGCAATTTACACCTGTGTGTGTATGTCAACACTGGT	1080
Qy	1081	GTCAATTGAAGCAATAGCACAGATCAGTGTCAAGACCTTACCCAAACCTCCAGGAACCTCT	1140
Db	1081	GTCAATTGAAGCAATAGCACAGATCAGTGTCAAGACCTTACCCAAACCTCCAGGAACCTCT	1140
Qy	1141	GTAGTGACCGAGAGCAGCTACAGCATCAAGCATCACTGACGTGGGACTCTGGGAACCTCTGAG	1200
Db	1141	GTAGTGACCGAGAGCAGCTACAGCATCAAGCATCACTGACGTGGGACTCTGGGAACCTCTGAG	1200
Qy	1201	CCTGTTTCTTTATTTACATAAATTCAGCATATAAACCCTTAAAGCTCTGAGGAACCTTTACAAAGAA	1260
Db	1201	CCTGTTTCTTTATTTACATAAATTCAGCATATAAACCCTTAAAGCTCTGAGGAACCTTTACAAAGAA	1260
Qy	1261	ATTGATGGGTGGGACCAACACGCTACAGTGTGCTGGACTTAAGTCCCTACTCGGATTTAT	1320
Db	1261	ATTGATGGGTGGGACCAACACGCTACAGTGTGCTGGACTTAAGTCCCTACTCGGATTTAT	1320
Qy	1321	GAATTCAGGGTGTGTGCTGTCAATAAATTTGGGGGGGGGCTCTCCAGCGAACCTCTGTGCTA	1380
Db	1321	GAATTCAGGGTGTGTGCTGTCAATAAATTTGGGGGGGGGCTCTCCAGCGAACCTCTGTGCTA	1380
Qy	1381	ACAAACCTCTCAGACGAGCAGCATCCAGTGTCCCGGAGGATGTCCAGGACGCAATGTG	1440
Db	1381	ACAAACCTCTCAGACGAGCAGCATCCAGTGTCCCGGAGGATGTCCAGGACGCAATGTG	1440
Qy	1441	AGTTCGACCAACCAATTTTGGTACAGTGGAGGAACCTGGAAGGCAAAATGGACAGATCCAA	1500
Db	1441	AGTTCGACCAACCAATTTTGGTACAGTGGAGGAACCTGGAAGGCAAAATGGACAGATCCAA	1500
Qy	1501	GGATATAGAGTTTATTTATACAAATGGATCCCACTCAACATGTCAACAACTGGATGAAACAC	1560
Db	1501	GGATATAGAGTTTATTTATACAAATGGATCCCACTCAACATGTCAACAACTGGATGAAACAC	1560
Qy	1561	AATGTAGCTGACAGCCAAATCACTACTATTGTGGCAACTTGTAGTCCCGGAGAAACATTTCT	1620
Db	1561	AATGTAGCTGACAGCCAAATCACTACTATTGTGGCAACTTGTAGTCCCGGAGAAACATTTCT	1620

Qy	1621	GTCAAAGTCTCGGCTTTTAACTCAATTGGAGATGGTCCCTTTCAAGTGAATACAAAGTC	1680
Db	1621	GTCAAAGTCTCGGCTTTTAACTCAATTGGAGATGGTCCCTTTCAAGTGAATACAAAGTC	1680
Qy	1681	ATCACTCAGACAGAGTACAGGCGAGCCACTAACTTCAAGCAGAACCTCGAGTCTGAA	1740
Db	1681	ATCACTCAGACAGAGTACAGGCGAGCCACTAACTTCAAGCAGAACCTCGAGTCTGAA	1740
Qy	1741	ACAAAGTATTTTGTCTCTTGGACACCTCCAGTTTCAGATACCAATTTGCCAACTATGAAGTC	1800
Db	1741	ACAAAGTATTTTGTCTCTTGGACACCTCCAGTTTCAGATACCAATTTGCCAACTATGAAGTC	1800
Qy	1801	GTCTACAAAGATGGGAGCATGGAGAGGCAACGAATTTACCATTTGAGCCAGGACATCA	1860
Db	1801	GTCTACAAAGATGGGAGCATGGAGAGGCAACGAATTTACCATTTGAGCCAGGAGACATCA	1860
Qy	1861	TATAGGCTGCAAGGACTGAAACCAACAGCTTATCTATTTCCGTCTGGCTGCAAGCTCC	1920
Db	1861	TATAGGCTGCAAGGACTGAAACCAACAGCTTATCTATTTCCGTCTGGCTGCAAGCTCC	1920
Qy	1921	CCTCAAGGCTGGGTGCTTCTACTGCAAGATATCAGTAGAACCATGCAAGTCAAGCGG	1980
Db	1921	CCTCAAGGCTGGGTGCTTCTACTGCAAGATATCAGTAGAACCATGCAAGTCAAGCGG	1980
Qy	1981	TCAGTCTCTCTCAAGACATTTAGTTGCAACGAGCCAAAGTTCCACTAGTATTTTGGTAAGT	2040
Db	1981	TCAGTCTCTCTCAAGACATTTAGTTGCAACGAGCCAAAGTTCCACTAGTATTTTGGTAAGT	2040
Qy	2041	TGGCAACCTCCACAGTGGGAAACAGAAATGGCAATTTCTTCCGACACT	2100
Db	2041	TGGCAACCTCCACAGTGGGAAACAGAAATGGCAATTTCTTCCGACACT	2100
Qy	2101	ACTCAGTGGATGGGAGATGCAAGCCTCAAGAGATTTTGGAAATTTCTTCCGACACT	2160
Db	2101	ACTCAGTGGATGGGAGATGCAAGCCTCAAGAGATTTTGGAAATTTCTTCCGACACT	2160
Qy	2161	ACCAAAATCTTTTGGAAACAGCTGGAAATTTGGAAATTTCTTCCGACACT	2220
Db	2161	ACCAAAATCTTTTGGAAACAGCTGGAAATTTGGAAATTTCTTCCGACACT	2220
Qy	2221	CATACAGATGTCGGCTCGGCTCGAGAGCTTGTCCGTGTTGATTCGAACCAATGAAGAT	2280
Db	2221	CATACAGATGTCGGCTCGGCTCGAGAGCTTGTCCGTGTTGATTCGAACCAATGAAGAT	2280
Qy	2281	GTTCTAGTGTCTCTCGCAAGTGGAGTAGAGCTGTCAAACCAATCTGTTAAA	2340
Db	2281	GTTCTAGTGTCTCTCGCAAGTGGAGTAGAGCTGTCAAACCAATCTGTTAAA	2340
Qy	2341	GTCTCATGGCGCTCAACCGTCCCAATAAACAGCATGGCCAGATTAAGAGGATATCAGGTG	2400
Db	2341	GTCTCATGGCGCTCAACCGTCCCAATAAACAGCATGGCCAGATTAAGAGGATATCAGGTG	2400
Qy	2401	CATTATGTGAGATGGGAAATGGTGAGCCCAAGGCGCAGCTGCTGAAAGATGTCAATG	2460
Db	2401	CATTATGTGAGATGGGAAATGGTGAGCCCAAGGCGCAGCTGCTGAAAGATGTCAATG	2460
Qy	2461	CTGGCTGATGACAGTGGGAAATTTGATGATCTACTGAACTGACATGATCATTTCTGGG	2520
Db	2461	CTGGCTGATGACAGTGGGAAATTTGATGATCTACTGAACTGACATGATCATTTCTGGG	2520
Qy	2521	CTCCAGCTGAAATCTTCTACTCCCTCACCGTCAAGGCTTACCAACCAAGGAGATGGT	2580
Db	2521	CTCCAGCTGAAATCTTCTACTCCCTCACCGTCAAGGCTTACCAACCAAGGAGATGGT	2580
Qy	2581	GCTCGACAGAGCCCAAACTGGTGTCCACCACTGGGCGAGTTCAGGGAACCTCGGCTT	2640
Db	2581	GCTCGACAGAGCCCAAACTGGTGTCCACCACTGGGCGAGTTCAGGGAACCTCGGCTT	2640
Qy	2641	GTGATTAACCACTCAGATGAATCTCTCTTATTTCAAGTGGCCACCTCCCGTGGACACA	2700
Db	2641	GTGATTAACCACTCAGATGAATCTCTCTTATTTCAAGTGGCCACCTCCCGTGGACACA	2700
Qy	2701	TTTGGACCTCTTCAGGCTACCGCTTAAATTTTGGCCGCAAGGATATGGAGCCACTTACT	2760
Db	2701	TTTGGACCTCTTCAGGCTACCGCTTAAATTTTGGCCGCAAGGATATGGAGCCACTTACT	2760
Qy	2761	ACTCTTGAAGTCTCTGAAAGAGAGATCACTTTTACAGCTACAGACATCCCAAGGGAGCA	2820
Db	2761	ACTCTTGAAGTCTCTGAAAGAGAGATCACTTTTACAGCTACAGACATCCCAAGGGAGCA	2820
Qy	2821	TCATACGTCTTTCAGGCTCTCAGCCAGAAACAAAGTGGGCTTTGGGGAGGAGATGGTGAAG	2880
Db	2821	TCATACGTCTTTCAGGCTCTCAGCCAGAAACAAAGTGGGCTTTGGGGAGGAGATGGTGAAG	2880
Qy	2881	GAGATTTCCATTTCCAGAGAGTACCAACTGGATTCCTTCAAAACCTTCTCACTCAGAAGGC	2940
Db	2881	GAGATTTCCATTTCCAGAGAGTACCAACTGGATTCCTTCAAAACCTTCTCACTCAGAAGGC	2940
Qy	2941	ACCACCTTCAACCTCCGTCAGTTTATCTTGGCAACCACTGTCTGGCAGAGAGAAATGGC	3000
Db	2941	ACCACCTTCAACCTCCGTCAGTTTATCTTGGCAACCACTGTCTGGCAGAGAGAAATGGC	3000
Qy	3001	ATTATCAACCAAGTATACCTTCTTTATAGGGATATCAACATCCCTTCTCCGATGGAG	3060
Db	3001	ATTATCAACCAAGTATACCTTCTTTATAGGGATATCAACATCCCTTCTCCGATGGAG	3060
Qy	3061	CAGCTTATTTGTTCCAGCTGACACCACTATGACACTCACTGGCTTAAACCAGATACACACA	3120
Db	3061	CAGCTTATTTGTTCCAGCTGACACCACTATGACACTCACTGGCTTAAACCAGATACACACA	3120
Qy	3121	TACGATGTTAAAGTACGTCTCATACGAGCAAGGGCCGGGCCATATAGTCCAGGTGC	3180
Db	3121	TACGATGTTAAAGTACGTCTCATACGAGCAAGGGCCGGGCCATATAGTCCAGGTGC	3180
Qy	3181	CAGTTCAAGGACATCGCTGTGGATCAAGTGTGTTGCAAAAAATTTTCAATGTCAGAGCAGTA	3240
Db	3181	CAGTTCAAGGACATCGCTGTGGATCAAGTGTGTTGCAAAAAATTTTCAATGTCAGAGCAGTA	3240
Qy	3241	ATCAAGACTTCCGCTGTGCTGTGGAGATTCAGAGAAATTTATTAATCTCCGCTGCCT	3300
Db	3241	ATCAAGACTTCCGCTGTGCTGTGGAGATTCAGAGAAATTTATTAATCTCCGCTGCCT	3300
Qy	3301	TTCAAAATCTTTTATGATGATGGGAAATGGTGAAGAGTGGATGGCCGAGCCACACAG	3360
Db	3301	TTCAAAATCTTTTATGATGATGGGAAATGGTGAAGAGTGGATGGCCGAGCCACACAG	3360
Qy	3361	AAGTTAAATGTCACCTGAAGCTGAGAAATCATATTTCTGCTGACAAATCTGTTGGA	3420
Db	3361	AAGTTAAATGTCACCTGAAGCTGAGAAATCATATTTCTGCTGACAAATCTGTTGGA	3420
Qy	3421	AACAGTCTGCTGGCTGACAGCAGGCTCACCGCAAGAGTGCACCAAGATGTATTACGT	3480
Db	3421	AACAGTCTGCTGGCTGACAGCAGGCTCACCGCAAGAGTGCACCAAGATGTATTACGT	3480
Qy	3481	ACCAAGCTCGCTTCAATTTGGGAGAGCAACTTGGATGGCATGATTTACTGTGCAACTGCCT	3540
Db	3481	ACCAAGCTCGCTTCAATTTGGGAGAGCAACTTGGATGGCATGATTTACTGTGCAACTGCCT	3540
Qy	3541	GAAGTACCTCAAAATGAGAAATATAAAAGGTTTACTACATAATATTTGTCCTTTCAAGAAA	3600
Db	3541	GAAGTACCTCAAAATGAGAAATATAAAAGGTTTACTACATAATATTTGTCCTTTCAAGAAA	3600
Qy	3601	TCTCGCGGAAATTTTATCAAGCCATGGGAGATGTCAGATGAAATAGGAATTTAGATGAGCTG	3660
Db	3601	TCTCGCGGAAATTTTATCAAGCCATGGGAGATGTCAGATGAAATAGGAATTTAGATGAGCTG	3660
Qy	3661	CTTAAAGGAGATATCTAGGAGCCGAGAAAGCATCGTTTATGGGAGAGAAATTTGAATTTAAG	3720
Db	3661	CTTAAAGGAGATATCTAGGAGCCGAGAAAGCATCGTTTATGGGAGAGAAATTTGAATTTAAG	3720
Qy	3721	CCATATATTGCGCTCACTTTGATGCTCTTCCCACTGAGTTTCACTTCCGCTGGGGATGACAAAG	3780
Db	3721	CCATATATTGCGCTCACTTTGATGCTCTTCCCACTGAGTTTCACTTCCGCTGGGGATGACAAAG	3780
Qy	3781	CATTATGTTGATTTTACAAACCAAGCAACTCCAAAGTGGTCAAGAAATATGTTCTTTTGTG	3840

Db 3781 CATTAAGTGGATTTTACAAACAGCACTCCAAAGTGGTCAAGAAATATGTCTTCTTTGTG 3840
Qy 3841 TTAGCAGTAATGGAAATGCGAGTCTAAGATGTATGCAACAGCAGCCCTTACTCCGACCCC 3900
Db 3841 TTAGCAGTAATGGAAATGCGAGTCTAAGATGTATGCAACAGCAGCCCTTACTCCGACCCC 3900
Qy 3901 GTGTGTCAATGGATCTGGATCCGAGCAATACGGATGAAGAAAGGCTTGAATCTGG 3960
Db 3901 GTGTGTCAATGGATCTGGATCCGAGCAATACGGATGAAGAAAGGCTTGAATCTGG 3960
Qy 3961 GTGTGTGTCTGTCTGTGAGTGTCTTTATCATCTGCTATGCTCATGCTATTTCTTCTT 4020
Db 3961 GTGTGTGTCTGTCTGTGAGTGTCTTTATCATCTGCTATGCTCATGCTATTTCTTCTT 4020
Qy 4021 TATAAAGGAAGAGGCGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAG 4080
Db 4021 TATAAAGGAAGAGGCGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAG 4080
Qy 4081 ATCCCTTCACACCAACCAAGCCTGTAGAACTGAGGCGCCTTAATCTTCAACACCG 4140
Db 4081 ATCCCTTCACACCAACCAAGCCTGTAGAACTGAGGCGCCTTAATCTTCAACACCG 4140
Qy 4141 GGTATGGCTAGCCATCTTCCAATACCCATCTTGGAACTTGCAGACCAATTTGAAGATTG 4200
Db 4141 GGTATGGCTAGCCATCTTCCAATACCCATCTTGGAACTTGCAGACCAATTTGAAGATTG 4200
Qy 4201 AAAGCAAAATGCAACTTGAAGTTTCCAGGAATATGAGTCAATTTGACCCCTGGCCAGCAG 4260
Db 4201 AAAGCAAAATGCAACTTGAAGTTTCCAGGAATATGAGTCAATTTGACCCCTGGCCAGCAG 4260
Qy 4261 TTCACTCTGGGAACATTTCAAACCTTGAAGTAAACAAACCAAGAAATAGATACCGCAATGTA 4320
Db 4261 TTCACTCTGGGAACATTTCAAACCTTGAAGTAAACAAACCAAGAAATAGATACCGCAATGTA 4320
Qy 4321 ATCCGATATGATCATTTCCGGGTTCTCTATCAGCTATAGAGGATCCAGGAGTGCAC 4380
Db 4321 ATCCGATATGATCATTTCCGGGTTCTCTATCAGCTATAGAGGATCCAGGAGTGCAC 4380
Qy 4381 TATGTGAATGCCAATCATAGATGGGTATAGGAAGCAAAATGCTTATTTGCAACACAG 4440
Db 4381 TATGTGAATGCCAATCATAGATGGGTATAGGAAGCAAAATGCTTATTTGCAACACAG 4440
Qy 4441 GGATCTCTCCCGGAAACATTTGGGACCTTTTGGAGAAATGATATGGGAACAAAGAGTGCC 4500
Db 4441 GGATCTCTCCCGGAAACATTTGGGACCTTTTGGAGAAATGATATGGGAACAAAGAGTGCC 4500
Qy 4501 ACAGTTGTATGATGACAAACTAGAAAGAAAGATCAAGGGTGAAGTGTGACCACTATTGG 4560
Db 4501 ACAGTTGTATGATGACAAACTAGAAAGAAAGATCAAGGGTGAAGTGTGACCACTATTGG 4560
Qy 4561 CCTAGCAGAGGCACAGAAACCCACGGACTCGTTCAAGTAAAGCTGTTGATPACTGTGGAG 4620
Db 4561 CCTAGCAGAGGCACAGAAACCCACGGACTCGTTCAAGTAAAGCTGTTGATPACTGTGGAG 4620
Qy 4621 CTGGCCACATATTGTGTTCGAACATTTGCACTTTTACAGAAATGTTCAAGTGAAGAGAGA 4680
Db 4621 CTGGCCACATATTGTGTTCGAACATTTGCACTTTTACAGAAATGTTCAAGTGAAGAGAGA 4680
Qy 4681 GAAGTGAGACAATTTCCAGTTTCCAGCTTCCGCTGATCATGTTTCCAGAACACCCCTACA 4740
Db 4681 GAAGTGAGACAATTTCCAGTTTCCAGCTTCCGCTGATCATGTTTCCAGAACACCCCTACA 4740
Qy 4741 CCTTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATG 4800
Db 4741 CCTTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATG 4800
Qy 4801 GTTGTGCACTGCACTGGGAGTTGGCGGACTGGTGTCTTTCATGCTCATAGATGCCATG 4860
Db 4801 GTTGTGCACTGCACTGGGAGTTGGCGGACTGGTGTCTTTCATGCTCATAGATGCCATG 4860
Qy 4861 TTAGAAAGAAATAGCAATGAAACCTGATATTTATGGCCATGTAATCTTTAATGAGA 4920
Db 4861 TTAGAAAGAAATAGCAATGAAACCTGATATTTATGGCCATGTAATCTTTAATGAGA 4920

Db 4861 TTAGAAAGAAATAGCAATGAAACCTGATATTTATGGCCATGTAATCTTTAATGAGA 4920
Qy 4921 GCCCAGAGGAACATATATGGTTTCAAAACAGAACCAATACATCTTTATCCATGATGCACTG 4980
Db 4921 GCCCAGAGGAACATATATGGTTTCAAAACAGAACCAATACATCTTTATCCATGATGCACTG 4980
Qy 4981 TTAGAACGAGTACTTGTGGAAATACCGAAGTGCAGCTAGAAAACCTTGTATGCTTACATTT 5040
Db 4981 TTAGAACGAGTACTTGTGGAAATACCGAAGTGCAGCTAGAAAACCTTGTATGCTTACATTT 5040
Qy 5041 CAGAAGCTGTACACAAATAGAAAAACGGGAGAGAAATGTACAGGAATGGAAGTTCGAATTTAAG 5100
Db 5041 CAGAAGCTGTACACAAATAGAAAAACGGGAGAGAAATGTACAGGAATGGAAGTTCGAATTTAAG 5100
Qy 5101 CGTCTAGCCAGCTCAAAAGCTCACCTCAAGGTTTATCAGTGCCTTCTTCCATGTAAT 5160
Db 5101 CGTCTAGCCAGCTCAAAAGCTCACCTCAAGGTTTATCAGTGCCTTCTTCCATGTAAT 5160
Qy 5161 AAATTTCAAAAATCGCCTTGTAAATATTTATGCCATATGAATCCCAAGGGTATGCTGCAG 5220
Db 5161 AAATTTCAAAAATCGCCTTGTAAATATTTATGCCATATGAATCCCAAGGGTATGCTGCAG 5220
Qy 5221 CCTATCCGTGGAGTAGAAGGATCTGATTACATCAATGCCAGTTTATTTATGATGGATACAGA 5280
Db 5221 CCTATCCGTGGAGTAGAAGGATCTGATTACATCAATGCCAGTTTATTTATGATGGATACAGA 5280
Qy 5281 CAACAGAAAGCCTACATCGCTACCCAGGGGCCCTTGGCAGAGCACCTGGAAGACTTCTGG 5340
Db 5281 CAACAGAAAGCCTACATCGCTACCCAGGGGCCCTTGGCAGAGCACCTGGAAGACTTCTGG 5340
Qy 5341 CGGATGCTCTGGGAACCAAAATTTCCACCATAGTTGTGATGCTCACCAGAGTGCCTGGAATG 5400
Db 5341 CGGATGCTCTGGGAACCAAAATTTCCACCATAGTTGTGATGCTCACCAGAGTGCCTGGAATG 5400
Qy 5401 GGCAGAGAGAAATGTCAACCAATACTGGCCAGAGAACGGTCTGCAAGATACCACTACTTTT 5460
Db 5401 GGCAGAGAGAAATGTCAACCAATACTGGCCAGAGAACGGTCTGCAAGATACCACTACTTTT 5460
Qy 5461 GTTGTAGATCCCATGGCTGAGTACAACATCCCAAGTATATCTTCAAGGAAATTTCAAGGTC 5520
Db 5461 GTTGTAGATCCCATGGCTGAGTACAACATCCCAAGTATATCTTCAAGGAAATTTCAAGGTC 5520
Qy 5521 ACAGATGCCAGGAGCGCCAGTCCCGAAGTAGTAAGGAGTTCAGTTTCACTGACTGCGCCA 5580
Db 5521 ACAGATGCCAGGAGCGCCAGTCCCGAAGTAGTAAGGAGTTCAGTTTCACTGACTGCGCCA 5580
Qy 5581 GAGCAAGGAGTGCCTGAGTCCGAGAGAGGATTTTATGACTTTCATCGGCCAAGTCCATAAA 5640
Db 5581 GAGCAAGGAGTGCCTGAGTCCGAGAGAGGATTTTATGACTTTCATCGGCCAAGTCCATAAA 5640
Qy 5641 ACAAAGAACAGTTTGGCCAAAGTAGGACCCATTTCACTGAGTCCATTTGAGGAGTT 5700
Db 5641 ACAAAGAACAGTTTGGCCAAAGTAGGACCCATTTCACTGAGTCCATTTGAGGAGTT 5700
Qy 5701 AGAATCTGGAGTCTTCAATACGCTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGGAGTT 5760
Db 5701 AGAATCTGGAGTCTTCAATACGCTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGGAGTT 5760
Qy 5761 GTAGATATCTTCCAGACTGTCAAAATGTTTAAAGAACCAACAGCAGCTATGTTGATACAGACA 5820
Db 5761 GTAGATATCTTCCAGACTGTCAAAATGTTTAAAGAACCAACAGCAGCTATGTTGATACAGACA 5820
Qy 5821 GAGGATCAATATCAGTTTCTTATCGTCCGCACTAGAGTACCTGGGAGCTTTTGACCCAC 5880
Db 5821 GAGGATCAATATCAGTTTCTTATCGTCCGCACTAGAGTACCTGGGAGCTTTTGACCCAC 5880
Qy 5881 TATGCAACGATAGAAAACCCCTGACCCATTTCTGGATTTTACTACAGGCCCTTCAATATCCA 5940
Db 5881 TATGCAACGATAGAAAACCCCTGACCCATTTCTGGATTTTACTACAGGCCCTTCAATATCCA 5940
Qy 5941 TGGAGTCTCTTCTGAGCCATACAGGGCCTTTGAGAGTCTTCTTAACTTCTAGCTAACA 6000
Db 5941 TGGAGTCTCTTCTGAGCCATACAGGGCCTTTGAGAGTCTTCTTAACTTCTAGCTAACA 6000

QY 1945 GCAGAAATATCAGCTAGAACCATGAGTCAAGAGCGTGAAGTCTCTCTCAAGACATTAATG 2004
DB 1443 GCAGAAATATCAGCTAGAACCATGAGTCAAGAGCGTGAAGTCTCTCTCAAGACATTAATG 1502
QY 2005 TGCACAGCCCAAGTTTCCACTAGTATTTTGGTAAAGTTGGCAACCTCCACAGTGAAAAA 2064
DB 1503 TGCACAGCCCAAGTTTCCACTAGTATTTTGGTAAAGTTGGCAACCTCCACAGTGAAAAA 1562
QY 2065 CAGAAATGGCAATATCACTGAATATCTCAATCAAGTACACTGCAAGTGGATGGGGAAGATGAC 2124
DB 1563 CAGAAATGGCAATATCACTGAATATCTCAATCAAGTACACTGCAAGTGGATGGGGAAGATGAC 1622
QY 2125 AAGCCTCAGGAGATTTTGGGAATTCCTTCGGACACTACCAATACTTTTGGAAACAGCTG 2184
DB 1623 AAGCCTCAGGAGATTTTGGGAATTCCTTCGGACACTACCAATACTTTTGGAAACAGCTG 1682
QY 2185 GAAAAATGGACTGAATATCCGAGTCACTGTGACAGCCCATACAGATGTGCGGCCCTGGCCCT 2244
DB 1683 GAAAAATGGACTGAATATCCGAGTCACTGTGACAGCCCATACAGATGTGCGGCCCTGGCCCT 1742
QY 2245 GAGAGCTTGTCCGTGTGATTGCAACCAATGAAGATGTTCTAGTGGTCTCTCCGCAAA 2304
DB 1743 GAGAGCTTGTCCGTGTGATTGCAACCAATGAAGATGTTCTAGTGGTCTCTCCGCAAA 1802
QY 2305 GTCGAGGTAGAGGCTGTCACTCAACATCTGTTAAAGTCTCATGCGGCTCACCCGTGCCC 2364
DB 1803 GTCGAGGTAGAGGCTGTCACTCAACATCTGTTAAAGTCTCATGCGGCTCACCCGTGCCC 1862
QY 2365 AATAAACAGCATGGCAGATGAAGATATCAGGTGCATATGTCAGGATGGAATGAT 2424
DB 1863 AATAAACAGCATGGCAGATGAAGATATCAGGTGCATATGTCAGGATGGAATGAT 1922
QY 2425 GAGCCAAAGGCGACGCCCATGCTGAAAGATGTCATGCTGGCTGCAAGTGCAGTGGGAATTT 2484
DB 1923 GAGCCAAAGGCGACGCCCATGCTGAAAGATGTCATGCTGGCTGCAAGTGCAGTGGGAATTT 1980
QY 2485 GATGATCTACTGAACATGACATGATCATTTCTGGGCTCCAGCTGGAACCTCTCTCTCC 2544
DB 1981 GATGAACT-----CATTTACATGATCATTTCTGGGCTCCAGCTGGAACCTCTCTCTCC 2034
QY 2545 CTCACGCTCAGACCTTACAAACAAAGAGATGTCGTCAGCAAGCCCAAACTGGTG 2604
DB 2035 CTCACGCTCAGACCTTACAAACAAAGAGATGTCGTCAGCAAGCCCAAACTGGTG 2094
QY 2605 TCCACCACTGGGGCAGTTCCAGGGAACCTCGGCTTGTGATTAAACCACTCAGATGAAT 2664
DB 2095 TCCACCACTGGGGCAGTTCCAGGGAACCTCGGCTTGTGATTAAACCACTCAGATGAAT 2154
QY 2665 ACTGCTCTTATCAGTGGCACCTCCGCTGGACACATTTGGACCTCTTCAGGGCTACCGT 2724
DB 2155 ACTGCTCTTATTCAGTGGCACCTCCGCTGGACACATTTGGACCTCTTCAGGGCTACCGT 2214
QY 2725 CTAAAAATTTGGCCGCAAGGATATGAGCGCACTTACTCTTTAGTTCTCTGAAAAAGAA 2784
DB 2215 CTAAAAATTTGGCCGCAAGGATATGAGCGCACTTACTCTTTAGTTCTCTGAAAAAGAA 2274
QY 2785 GATCACTTTACAGCTACAGACATCCAAAGGAGCATCATAGTCTTCAGGCTCTCAGCC 2844
DB 2275 GATCACTTTACAGCTACAGACATCCAAAGGAGCATCATAGTCTTCAGGCTCTCAGCC 2334
QY 2845 AGAAACAAAGTGGGCTTTGGGAGGAGATGTTGAAGGAGATTTCCATTCAGAGAAAGTA 2904
DB 2335 AGAAACAAAGTGGGCTTTGGGAGGAGATGTTGAAGGAGATTTCCATTCAGAGAAAGTA 2394
QY 2905 CCAACTGGATTTCCCTCAAAACCTTCACTCAGAAGGCCACCTTCAACCTCCGCTCCAGTTA 2964
DB 2395 CCAACTGGATTTCCCTCAAAACCTTCACTCAGAAGGCCACCTTCAACCTCCGCTCCAGTTA 2454
QY 2965 TCTTGGCAACCACTGCTCGCAGAGAGAAATGGCAATTTATCACCAGTATACCCCTTCTT 3024
DB 2455 TCTTGGCAACCACTGCTCGCAGAGAGAAATGGCAATTTATCACCAGTATACCCCTTCTT 2514

QY 3025 TATAGGAGTATCAACATCCCCCTTCTCCGATGGAGCAGCTTATTTGTTCCAGCTGACACC 3084
DB 2515 TATAGGAGTATCAACATCCCCCTTCTCCGATGGAGCAGCTTATTTGTTCCAGCTGACACC 2574
QY 3085 ACTATGACACTCACTGGCTTAAAAACAGATACCATACGATGTAAAAAGTACGCTCAT 3144
DB 2575 ACTATGACACTCACTGGCTTAAAAACAGATACCATACGATGTAAAAAGTACGCTCAT 2634
QY 3145 ACAGAGAAAGGCGCGGCCCATATAGTCCAGTGTCCAGTTCAGGACACTGCTGTGGAT 3204
DB 2635 ACAGAGAAAGGCGCGGCCCATATAGTCCAGTGTCCAGTTCAGGACACTGCTGTGGAT 2694
QY 3205 CAAGTGTTCGAAAAATTTTTCATGTCAAGCAGTAAATGAAGACTTCGCTGTGCTGTCT 3264
DB 2695 CAAGTGTTCGAAAAATTTTTCATGTCAAGCAGTAAATGAAGACTTCGCTGTGCTGTCT 2754
QY 3265 TGGGAGATTCAGAGAAATTTAACTCCGCCATGCTTTCAAAAATTTCTTTATGATGATGGG 3324
DB 2755 TGGGAGATTCAGAGAAATTTAACTCCGCCATGCTTTCAAAAATTTCTTTATGATGATGGG 2814
QY 3325 AAAATGTTAGAAAGTGGATGGCGGAGCCACAG - AAGTTAAATGTTCAACCTGAAGCC 3383
DB 2815 AAAATGTTAGAAAGTGGATGGCGGAGCCACAGAAAAGTTAAATGTTCAACCTGAAGCC 2874
QY 3384 TGAGAAATCATATTTTCATTTGCTGACAAATCGTGGAAACAGTCTGCTGGCTGCAGCA 3443
DB 2875 TGAGAAATCATATTTTCATTTGCTGACAAATCGTGGAAACAGTCTGCTGGCTGCAGCA 2934
QY 3444 CAGGCTCAGGCAAGACTGCACAGATGTATTACGTACCAAGCTGCTCTTCAATGGGAA 3503
DB 2935 CAGGCTCAGGCAAGACTGCACAGATGTATTACGTACCAAGCTGCTCTTCAATGGGAA 2994
QY 3504 GACCAACTGGATGGCATGATTCTGTGCAACTGCTGAAAGTACCTGCAAAATGAGAAATAT 3563
DB 2995 GACCAACTGGATGGCATGATTCTGTGCAACTGCTGAAAGTACCTGCAAAATGAGAAATAT 3054
QY 3564 AAAAGGTTACTACATAATTTGTCCTTTGAAGAAATCTCGCGGAAATTTTCAAGCC 3623
DB 3055 AAAAGGTTACTACATAATTTGTCCTTTGAAGAAATCTCGCGGAAATTTTCAAGCC 3114
QY 3624 ATGGGAGATCCAGATGAAATTTAGATGAGTCTGCTTAAAGAGATATCTAGGAAGCG 3683
DB 3115 ATGGGAGATCCAGATGAAATTTAGATGAGTCTGCTTAAAGAGATATCTAGGAAGCG 3174
QY 3684 CAGAAGCATCCGTTATGGGAGAGAAAGTTGAATTTAAAGCCATATATGCGGCTCACTTTGA 3743
DB 3175 CAGAAGCATCCGTTATGGGAGAGAAAGTTGAATTTAAAGCCATATATGCGGCTCACTTTGA 3234
QY 3744 TGTCTTTCCCACTGAGTTCACTGGGGGATGACAGCATTTATGCTGATTTTCAAAACAA 3803
DB 3235 TGTCTTTCCCACTGAGTTCACTGGGGGATGACAGCATTTATGCTGATTTTCAAAACAA 3294
QY 3804 GCAACTCAAAAGTGGTCAAGAAATATGCTCTTTGTTAGCAGTAAATGGAACATGCAGA 3863
DB 3295 GCAACTCAAAAGTGGTCAAGAAATATGCTCTTTGTTAGCAGTAAATGGAACATGCAGA 3354
QY 3864 GTCTAAGATGTATGCAACCAAGCCCTTACTCCGACCCCGTGGTGTCAATGGAATCTGGATCC 3923
DB 3355 GTCTAAGATGTATGCAACCAAGCCCTTACTCCGACCCCGTGGTGTCAATGGAATCTGGATCC 3414
QY 3924 GCAGCCAAATCAGGATCAAGAGAGGCTTGTAGTGGTGTAGGCTCTGCTTGCAGT 3983
DB 3415 GCAGCCAAATCAGGATCAAGAGAGGCTTGTAGTGGTGTAGGCTCTGCTTGCAGT 3474
QY 3984 GGTCTTTTATCATCTGCAATTTGCTATTGCTTTTATTTTAAAAAGGAGAGGCGCAGAGTC 4043
DB 3475 GGTCTTTTATCATCTGCAATTTGCTATTGCTTTTATTTTAAAAAGGAGAGGCGCAGAGTC 3534
QY 4044 CGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAGATCCCTTTCACACCAACCAACAGA 4103
DB 3535 CGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAGATCCCTTTCACACCAACCAACAGA 3594
QY 4104 CCCTGTAGAACTGAGGCGCCCTTAACTTTCAAAACACCGGGTATGCTAGCCATCTCCAT 4163

3595	Db		CCCTGTAGAACTGAGGGCGCCTTAACTTTCAACAACCGGGTATGGCTAGCATCCTCCAAT	3654
4164	Qy		ACCCATCTTGGAACTTCGACAGCCACATTTGAAAGATTGAAAGCAAAATGACAACCTTGAAGTT	4223
3655	Db		ACCCATCTTGGAACTTCGACAGCCACATTTGAAAGATTGAAAGCAAAATGACAACCTTGAAGTT	3714
4224	Qy		TTTCCAGGAATATGAGTCAATTGACCTTGGCCAGCAGTTCACTTGGGAAACATTCAAACTT	4283
3715	Db		TTTCCAGGAATATGAGTCAATTGACCTTGGCCAGCAGTTCACTTGGGAAACATTCAAACTT	3774
4284	Qy		GGAAGTAAACAAACCAAAGAATAGATACGGGAATGTAAATCGCATATGATCATTTCCCGGGT	4343
3775	Db		GGAAGTAAACAAACCAAAGAATAGATACGGGAATGTAAATCGCATATGATCATTTCCCGGGT	3834
4344	Qy		TCTCTATCAGCTATAGNAGGATCCAGAGAGTGACTATGTGAATGCCAACTACATAGA	4403
3835	Db		TCTCTATCAGCTATAGNAGGATCCAGAGAGTGACTATGTGAATGCCAACTACATAGA	3894
4404	Qy		TGGGTATAGGAAGCAAAATGCGTATATTGCAACACAGGGATCTCTCCCGAAACATTTGG	4463
3895	Db		TGGGTATAGGAAGCAAAATGCGTATATTGCAACACAGGGATCTCTCCCGAAACATTTGG	3954
4464	Qy		GGACTTTTGGAGAAATGATATGGGAACAAACGGAGTGCACAGTTGTGTATGATGACAAACT	4523
3955	Db		GGACTTTTGGAGAAATGATATGGGAACAAACGGAGTGCACAGTTGTGTATGATGACAAACT	4014
4524	Qy		AGAAGAAAGATCAAGGGTGAAGTGTGACAGTATTTGGCCTAGCAGAGGCACAGAAACCCA	4583
4015	Db		AGAAGAAAGATCAAGGGTGAAGTGTGACAGTATTTGGCCTAGCAGAGGCACAGAAACCCA	4074
4584	Qy		CGGACTCGTTCAAGTAAACGCTGCTTGATCTGCTGGAGCTGGCCACATATTGTGTTCGAAC	4643
4075	Db		CGGACTCGTTCAAGTAAACGCTGCTTGATCTGCTGGAGCTGGCCACATATTGTGTTCGAAC	4134
4644	Qy		ATTTGCACCTTTACAAGAAATGGTTCAAGTGAGAAGAGAAAGTGAGACAAATTCAGTTTCAC	4703
4135	Db		ATTTGCACCTTTACAAGAAATGGTTCAAGTGAGAAGAGAAAGTGAGACAAATTCAGTTTCAC	4194
4704	Qy		CGCCTGGCTGATCATTGTGTTCAGAAACACCCCTACACCTTTTCTAGCTTCTTACGTAG	4763
4195	Db		CGCCTGGCTGATCATTGTGTTCAGAAACACCCCTACACCTTTTCTAGCTTCTTACGTAG	4254
4764	Qy		AGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATGGTTGTGCACCTGCAGTGGCGAGT	4823
4255	Db		AGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATGGTTGTGCACCTGCAGTGGCGAGT	4314
4824	Qy		TGGCCGGACTGGTTGCTTCATCGTCAATAGATGCCATGTTAGAAAGAAATAAGACATGAAAA	4883
4315	Db		TGGCCGGACTGGTTGCTTCATCGTCAATAGATGCCATGTTAGAAAGAAATAAGACATGAAAA	4374
4884	Qy		AACGTGTAGATATTTATGGCCCATGTAACTTTTAATCAGAGCCACAGGAACTATATGGTTCA	4943
4375	Db		AACGTGTAGATATTTATGGCCCATGTAACTTTTAATCAGAGCCACAGGAACTATATGGTTCA	4434
4944	Qy		AACAGAGACCAATACATCTTTTATCCATGATGCCATGTTTAAAGCAGTGACTCTGTGAAA	5003
4435	Db		AACAGAGACCAATACATCTTTTATCCATGATGCCATGTTTAAAGCAGTGACTCTGTGAAA	4494
5004	Qy		TACCGAAGTGCAGCTAGAAAACCTTTGTATGCCCTACATTCAGAAAGCTGACACAAATAGAAAC	5063
4495	Db		TACCGAAGTGCAGCTAGAAAACCTTTGTATGCCCTACATTCAGAAAGCTGACACAAATAGAAAC	4554
5064	Qy		GGGAGAGAAATGTCACAGGAATGGAGCTCGAATTTTAAGGCTTAGCCAGCTCCTAAAGCTCA	5123
4555	Db		GGGAGAGAAATGTCACAGGAATGGAGCTCGAATTTTAAGGCTTAGCCAGCTCCTAAAGCTCA	4614
5124	Qy		CACCTCAAGGTTTATCAGTGGCAATCTTCCATGTAAATAAATCAAAAATCGCCTTGTATA	5183
4615	Db		CACCTCAAGGTTTATCAGTGGCAATCTTCCATGTAAATAAATCAAAAATCGCCTTGTATA	4674
5184	Qy		TATTATGCCCATATGAATCCAAAGGGTATGGCTCGAGCCTATCCGTGGAGTAGAAGGATC	5243

4675	TATTATGCCATATGAATCCAAAGGGGTATGCCCTGCAGCCTATCCGTGGAGTAGAAGATC	4733
5244	TGATTTACATCAATGCCAGTTTATTTGATGGATACAGACAACAGAAAGCCTACATCGCTAC	5303
4735	TGATTACATCAATGCCAGTTTATTTGATGGATACAGACAACAGAAAGCCTACATCGCTAC	4794
5304	CCAGGGGCCCTTGGCAGAGACCACTGAAGAATCTTGGCCGATGCTCTGGGAACACAATTC	5363
4795	CCAGGGGCCCTTGGCAGAGACCACTGAAGAATCTTGGCCGATGCTCTGGGAACACAATTC	4854
5364	CACCATAGTTGTGATGCTCTCCAAAGCTGCGTGAATTCGGGCAGAGAGAAATGTCACCAATA	5423
4855	CACCATAGTTGTGATGCTCTCCAAAGCTGCGTGAATTCGGGCAGAGAGAAATGTCACCAATA	4914
5424	CTGGCCAGCAGAACGGTCTCCAAAGATACCAAGTACTTTGTTGTAGATCCCATGGCTCAGTA	5483
4915	CTGGCCAGCAGAACGGTCTCCAAAGATACCAAGTACTTTGTTGTAGATCCCATGGCTCAGTA	4974
5484	CAACATGCCACAGTATATCTTAAGGGAAATTCAAAGGTCACAGATGCCAGGACGGCCAGTC	5543
4975	CAACATGCCACAGTATATCTTAAGGGAAATTCAAAGGTCACAGATGCCAGGACGGCCAGTC	5034
5544	CCGACAGTAAAGCAGTCCAGTCTCACTGACTGCGCCAGACCAAGGAGTCCCAAGTCCGG	5603
5035	CCGACAGTAAAGCAGTCCAGTCTCACTGACTGCGCCAGACCAAGGAGTCCCAAGTCCGG	5094
5604	AGAAGGATTTATTGACTTTCATTCGGCCAAAGTCCATAAAACAAAAGAACAGTTTGGCCAAAG	5663
5095	AGAAGGATTTATTGACTTTCATTCGGCCAAAGTCCATAAAACAAAAGAACAGTTTGGCCAAAG	5154
5664	TGGACCCATTTCAAGTCCATTTGCAGCGCGCGTTTGGAAAGAACTGGAGTCTTCATAAGCCT	5723
5155	TGGACCCATTTCAAGTCCATTTGCAGCGCGCGTTTGGAAAGAACTGGAGTCTTCATAAGCCT	5214
5724	AAGCATTTGTTTGGAAAGATGAGATATGAAGGAGTTGTAGATATCTTCAGAGCTGTCAA	5783
5215	AAGCATTTGTTTGGAAAGATGAGATATGAAGGAGTTGTAGATATCTTCAGAGCTGTCAA	5274
5784	AATGTTAAGAACACAAACGACAGCTATGGTACAGACAGAGGATCAATATCAGTTTTTCCTA	5843
5275	AATGTTAAGAACACAAACGACAGCTATGGTACAGACAGAGGATCAATATCAGTTTTTCCTA	5334
5844	TCGTGCCGCACTAGAGTACTTGGCAGCTTTGACCACTATGCAACGTAGAAAACCCCTGAC	5903
5335	TCGTGCCGCACTAGAGTACTTGGCAGCTTTGACCACTATGCAACGTAGAAAACCCCTGAC	5394
5904	CCATTCTGGATTTTACTACAGGCCCTTCAATATCCATGGAGTCTCTCTGAGCCATACA	5963
5395	CCATTCTGGATTTTACTACAGGCCCTTCAATATCCATGGAGTCTCTCTGAGCCATACA	5454
5964	GGGCACTTGAGAAGTCCCTTTAACTTCTAGCTAAACAATCTACTAGTGGGACTATTACAC	6023
5455	GGGCACTTGAGAAGTCCCTTTAACTTCTAGCTAAACAATCTACTAGTGGGACTATTACAC	5514
6024	ACAAAACAAAATTAATAAATAATTTCCAGGTGGACCAAGAAATCTTTTGACATCGCCCT	6083
5515	ACAAAACAAAATTAATAAATAATTTCCAGGTGGACCAAGAAATCTTTTGACATCGCCCT	5574
6084	TCCACACATCTGCTCATATAACATTTTAGGGCCAAAGGGAGGGAATGTTTAAAAAGA	6143
5575	TCCACACATCTGCTCATATAACATTTTAGGGCCAAAGGGAGGGAATGTTTAAAAAGA	5634
6144	AAGTCTTTGATTTAGTTTTTTTAGTTATTGTTAAAGATATCTGCTGACTCTGCTTCAATTTCTA	6203
5635	AAGTCTTTGATTTAGTTTTTTTAGTTATTGTTAAAGATATCTGCTGACTCTGCTTCAATTTCTA	5694
6204	ACTGTGTAACCTTTTTTTTAAACAAATGTATCATTCGATAAAGTGAAATTTTAAAAAGTT	6263
5695	ACTGTGTAACCTTTTTTTTAAACAAATGTATCATTCGATAAAGTGAAATTTTAAAAAGTT	5754
Db	RESULT	4
Qy	HSPTPD	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2005, 22:23:57 ; Search time 923 Seconds
(without alignments)
11102.926 Million cell updates/sec

Title: US-10-772-636-63

Perfect score: 6263

Sequence: 1 gctaaccaaggagagcgc.....aagtgaatttttaaaaaagttt 6263

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6263	100.0	6263	4	US-09-023-655-1473
2	2473	39.5	6363	4	Sequence 1473, Ap
3	2398	38.3	6000	1	Sequence 879, App
4	2398	38.3	6000	2	Sequence 6, Appli
5	2398	38.3	6000	3	Sequence 6, Appli
6	2398	38.3	6000	5	Sequence 6, Appli
7	2357.4	37.6	7702	4	Sequence 1336, Ap
8	2357.4	37.6	7702	4	Sequence 1336, Ap
9	1345.8	21.5	4078	4	Sequence 1132, Ap
10	1300	20.8	5690	2	Sequence 2, Appli
11	1300	20.8	5690	2	Sequence 2, Appli
12	1095.8	17.5	3467	4	Sequence 1, Appli
13	538	8.6	1029	4	Sequence 23, Appli
14	538	8.6	1029	4	Sequence 24, Appli
15	442	7.1	2872	1	Sequence 4, Appli
16	442	7.1	2872	4	Sequence 4, Appli
17	438.8	7.0	2409	1	Sequence 2, Appli
18	438.8	7.0	2409	4	Sequence 2, Appli
19	438.8	7.0	2537	4	Sequence 1485, Ap
20	438.8	7.0	2565	4	Sequence 1311, Ap
21	438.8	7.0	3615	4	Sequence 1485, Ap
22	376.8	6.0	1689	4	Sequence 1254, Ap
23	348	5.6	4651	2	Sequence 4915, Ap
24	348	5.6	4651	2	Sequence 4, Appli
25	348	5.6	4651	2	Sequence 4, Appli
26	338.4	5.4	4374	2	Sequence 945, App
27	338.4	5.4	4374	2	Sequence 3, Appli

28 321.4 5.1 4078 4 US-09-016-434-1134 Sequence 1134, Ap
29 317.2 5.1 3228 4 US-09-949-016-2437 Sequence 2437, Ap
30 317.2 5.1 4998 4 US-09-023-655-1299 Sequence 1299, Ap
31 270 4.3 5769 1 US-08-652-971-1 Sequence 1, Appli
32 270 4.3 5769 2 US-08-991-258A-1 Sequence 1, Appli
33 270 4.3 5769 2 US-08-769-399-1 Sequence 1, Appli
34 270 4.3 5769 3 US-08-991-953A-1 Sequence 1, Appli
35 267.4 4.3 2422 4 US-09-949-016-1287 Sequence 1287, Ap
36 265 4.2 7941 3 US-09-816-703A-1 Sequence 1, Appli
37 260.2 4.2 6924 1 US-08-015-973-2 Sequence 2, Appli
38 260.2 4.2 6924 2 US-08-448-164-2 Sequence 2, Appli
39 259.6 4.1 3996 4 US-09-949-016-2026 Sequence 2026, Ap
40 259.6 4.1 4338 1 US-08-015-986A-1 Sequence 1, Appli
41 259.6 4.1 4338 2 US-08-446-363-1 Sequence 1, Appli
42 257.8 4.1 4389 4 US-09-023-655-991 Sequence 991, App
43 256.6 4.1 5581 4 US-09-023-655-966 Sequence 966, App
44 220.2 3.5 4798 4 US-09-620-312D-298 Sequence 298, App
45 218.8 3.5 2520 4 US-09-949-016-4698 Sequence 4698, Ap

ALIGNMENTS

RESULT 1

US-09-023-655-1473
; Sequence 1473, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1473:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6263 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g755652
US-09-023-655-1473

Query Match

100.0%; Score 6263; DB 4; Length 6263;

Best Local Similarity 100.0%; Pred. No. 0; Matches 6263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GCTAACTCAAGGAGAGAGCTCTGTGTGAACACCCGCTGGGATCTAAAGAAACAAGCTCTGAAG	60
Db	1	GCTAACTCAAGGAGAGAGCTCTGTGTGAACACCCGCTGGGATCTAAAGAAACAAGCTCTGAAG	60
Qy	61	TGTTCCAGCTGAATTTTACAGATCGGACAGACTCGCTGGGCTCCGGAGGAGGATGATTCGA	120
Db	61	TGTTCCAGCTGAATTTTACAGATCGGACAGACTCGCTGGGCTCCGGAGGAGGATGATTCGA	120
Qy	121	AGCTGCTCGGCGACGCTGCTCCAGCTGCGAGGATGTCAGCTAGCCAGGCTGCTGCTG	180
Db	121	AGCTGCTCGGCGACGCTGCTCCAGCTGCGAGGATGTCAGCTAGCCAGGCTGCTGCTG	180
Qy	181	CTGCTCTCTCACTTTCTCTCCGCGACGAGTCTCAGACACCTCCAAAGGTTTACACGAACA	240
Db	181	CTGCTCTCTCACTTTCTCTCCGCGACGAGTCTCAGACACCTCCAAAGGTTTACACGAACA	240
Qy	241	CCGTTGATCAGACAGGGGCTCTGGCGGAGTTGCTCTCTTTCATCTCGCAAGCTACGGGA	300
Db	241	CCGTTGATCAGACAGGGGCTCTGGCGGAGTTGCTCTCTTTCATCTCGCAAGCTACGGGA	300
Qy	301	GACCCAAAGCTTAAATTTGCTGGAACAAAGAAAGAAAGTCAAGATTCAGAGATTT	360
Db	301	GACCCAAAGCTTAAATTTGCTGGAACAAAGAAAGAAAGTCAAGATTCAGAGATTT	360
Qy	361	GAGTAATAGAGTTTACAGATGGGCTCTGGATCAGTTCTCAGAATACAACCTTACGGACT	420
Db	361	GAGTAATAGAGTTTACAGATGGGCTCTGGATCAGTTCTCAGAATACAACCTTACGGACT	420
Qy	421	CCGAGGATGAGGCCATTTATGAATGTGGCCCTCAATATATGTGGGAGAAATTAAGTGA	480
Db	421	CCGAGGATGAGGCCATTTATGAATGTGGCCCTCAATATATGTGGGAGAAATTAAGTGA	480
Qy	481	TCCACACAGCTCAGATTTTGGCGGAGATCAAAATCCAGGGGCTTCCCTTACCATTTGAC	540
Db	481	TCCACACAGCTCAGATTTTGGCGGAGATCAAAATCCAGGGGCTTCCCTTACCATTTGAC	540
Qy	541	ATGGGCCCAAGTTGAAAGTGTGTGAGGTACTCGCACGGCCACCATGCTTTGTGCAGCC	600
Db	541	ATGGGCCCAAGTTGAAAGTGTGTGAGGTACTCGCACGGCCACCATGCTTTGTGCAGCC	600
Qy	601	AGTGGTAATCCGGATCAGAAATCACTTGTGTTTAAAGATTTTATATGTGCAGAGCTCGGA	660
Db	601	AGTGGTAATCCGGATCAGAAATCACTTGTGTTTAAAGATTTTATATGTGCAGAGCTCGGA	660
Qy	661	AACAAATGCTGCTTATTAAGCAGTTACGATCAGAAATCTATTGTTGTTACACCAATGA	720
Db	661	AACAAATGCTGCTTATTAAGCAGTTACGATCAGAAATCTATTGTTGTTACACCAATGA	720
Qy	721	GGAGCCCTTCAGATTTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGAGTGTGTTGCC	780
Db	721	GGAGCCCTTCAGATTTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGAGTGTGTTGCC	780
Qy	781	ACCAACAGCGGGGCACTCGCTATTTCGCTCTCCGCAATTTATATGTGCAGAGCTCGGA	840
Db	781	ACCAACAGCGGGGCACTCGCTATTTCGCTCTCCGCAATTTATATGTGCAGAGCTCGGA	840
Qy	841	GAGTTGCGCTGTCACCAAGATTTCTATCCACCCCTCAATCATGAATCATGCCA	900
Db	841	GAGTTGCGCTGTCACCAAGATTTCTATCCACCCCTCAATCATGAATCATGCCA	900
Qy	901	GGCGGAAGCTGTAATATACCTGTGTGGCGGCTGAGGATGATGCCAATGGAAGATGTCGA	960
Db	901	GGCGGAAGCTGTAATATACCTGTGTGGCGGCTGAGGATGATGCCAATGGAAGATGTCGA	960
Qy	961	ATGTTGGGGGAGAGATCTGACACCTGGAAGATGATGCCAATGGAAGATGTCGA	1020
Db	961	ATGTTGGGGGAGAGATCTGACACCTGGAAGATGATGCCAATGGAAGATGTCGA	1020
Qy	1021	GAACTGAATGATGTAAGACAGTCAAGCAATTTACACCTGTGTTGCTATGTCAACACTGGGT	1080
Db	1021	GAACTGAATGATGTAAGACAGTCAAGCAATTTACACCTGTGTTGCTATGTCAACACTGGGT	1080
Qy	1081	GTCAATTGAAGCAATAGCACAGATCACTGTCAAAGCTTTACCCAAACCTCCAGGAACTCCT	1140
Db	1081	GTCAATTGAAGCAATAGCACAGATCACTGTCAAAGCTTTACCCAAACCTCCAGGAACTCCT	1140
Qy	1141	GTAGTGACCGAGAGCAGCTCAAGCATCACTGACGTGGGACTCTCGGAAACCTCTGAG	1200
Db	1141	GTAGTGACCGAGAGCAGCTCAAGCATCACTGACGTGGGACTCTCGGAAACCTCTGAG	1200
Qy	1201	CTGTGTTCTTATTACATAAATTCAGCATAAACCTTAAAGAACTCTGAGGAACTTTACAAAGAA	1260
Db	1201	CTGTGTTCTTATTACATAAATTCAGCATAAACCTTAAAGAACTCTGAGGAACTTTACAAAGAA	1260
Qy	1261	ATTGATGGGTGGCGACCAACACGCTACAGTGTGCTGGACTAAGTCTCTACTCGGATAT	1320
Db	1261	ATTGATGGGTGGCGACCAACACGCTACAGTGTGCTGGACTAAGTCTCTACTCGGATAT	1320
Qy	1321	GAATTCAGGGTTCCTGCTGCTCAATTAACATTTGGGGGGGGCTCCACAGCAACCTGTGCTA	1380
Db	1321	GAATTCAGGGTTCCTGCTGCTCAATTAACATTTGGGGGGGGCTCCACAGCAACCTGTGCTA	1380
Qy	1381	ACAAACCTTCAGAGCAAGCAACATCCAGTGTGCTGGAGGATGTCCAGGCAACGATGTTG	1440
Db	1381	ACAAACCTTCAGAGCAAGCAACATCCAGTGTGCTGGAGGATGTCCAGGCAACGATGTTG	1440
Qy	1441	AGTTGACACCACTTTTGGTACAGTGGAGAAAGCTTGAAGGCCAAATGACAGATCCAA	1500
Db	1441	AGTTGACACCACTTTTGGTACAGTGGAGAAAGCTTGAAGGCCAAATGACAGATCCAA	1500
Qy	1501	GGATATAGAGTTTATTAACAATGGATCCCACTCAACATGTCAACACTGATGAAACAC	1560
Db	1501	GGATATAGAGTTTATTAACAATGGATCCCACTCAACATGTCAACACTGATGAAACAC	1560
Qy	1561	AATGTAGCTCAGACCAAAATCACTACTATTGGCAACTTAGTGCCCAAGAAACATATTTCT	1620
Db	1561	AATGTAGCTCAGACCAAAATCACTACTATTGGCAACTTAGTGCCCAAGAAACATATTTCT	1620
Qy	1621	GTCAAAGTCTGGCTTTTACTCAATTTGGAGATGGTCCCTTTCAAGTGACATCAAGTGC	1680
Db	1621	GTCAAAGTCTGGCTTTTACTCAATTTGGAGATGGTCCCTTTCAAGTGACATCAAGTGC	1680
Qy	1681	ATCACTCAGACAGAGTACAGGCGCAGCTTAACTTCAAGCAGAACTTGAGTCTGAA	1740
Db	1681	ATCACTCAGACAGAGTACAGGCGCAGCTTAACTTCAAGCAGAACTTGAGTCTGAA	1740
Qy	1741	ACAAGTATTTTGTCTCTTGGACACCTTCAAGTTCAGATACCAATTTGCAACTGAACTG	1800
Db	1741	ACAAGTATTTTGTCTCTTGGACACCTTCAAGTTCAGATACCAATTTGCAACTGAACTG	1800
Qy	1801	GTCTAACAAGATGGGAGCATGGAGAGGAGCAAGAAATTAACATTGAGCCAGGAGACATCA	1860
Db	1801	GTCTAACAAGATGGGAGCATGGAGAGGAGCAAGAAATTAACATTGAGCCAGGAGACATCA	1860
Qy	1861	TATAGGCTGCAAGSACTGAAACCAACAGCTTACTATTTCCTGCTGGCTGACGCTCC	1920
Db	1861	TATAGGCTGCAAGSACTGAAACCAACAGCTTACTATTTCCTGCTGGCTGACGCTCC	1920
Qy	1921	CCTCAAGGCTGGGTGCTTCTTACTGCAAGAAATATCAGCTAGAACCATGCAAGTCAAGCG	1980
Db	1921	CCTCAAGGCTGGGTGCTTCTTACTGCAAGAAATATCAGCTAGAACCATGCAAGTCAAGCG	1980
Qy	1981	TCAGTCTCTCTCAAGACATTTAGTTGCAACAGCCCAAGTTCCACTAGTATTTTGGTAAGT	2040
Db	1981	TCAGTCTCTCTCAAGACATTTAGTTGCAACAGCCCAAGTTCCACTAGTATTTTGGTAAGT	2040
Qy	2041	TGCAACCTCCACAGTGGAAACAGAAATGGGATTAATCACTGAATATCTCCATCAAGTAC	2100
Db	2041	TGCAACCTCCACAGTGGAAACAGAAATGGGATTAATCACTGAATATCTCCATCAAGTAC	2100
Qy	2101	ACTGCAAGTGGGAGAGATGACAGCTTACAGATTTTGGGAAATTTCTTCCGACACT	2160
Db	2101	ACTGCAAGTGGGAGAGATGACAGCTTACAGATTTTGGGAAATTTCTTCCGACACT	2160

QY 2161 ACCAAATACCTTTTGGAAACAGCTGGAAATAATGGAATGGAATACCGGATCACTGTGACAGCC 2220
DB 2161 ACCAAATACCTTTTGGAAACAGCTGGAAATAATGGAATGGAATACCGGATCACTGTGACAGCC 2220
QY 2221 CATACAGATGTCGGCCCTCGCCCTCAGAGCTTGTCCGTGTTGATTCGAACCAATCAAGAT 2280
DB 2221 CATACAGATGTCGGCCCTCGCCCTCAGAGCTTGTCCGTGTTGATTCGAACCAATCAAGAT 2280
QY 2281 GTTCTAGTGTCCTCTCGCAAGTTCGAGGTAGAGGCTGTCAACTCAACATCTCTGTAAA 2340
DB 2281 GTTCTAGTGTCCTCTCGCAAGTTCGAGGTAGAGGCTGTCAACTCAACATCTCTGTAAA 2340
QY 2341 GTCTCATGCGCTCAACCGTGGCCCAATAAACAGCATGCGCCAGATTAAGAGGATATCAGGTG 2400
DB 2341 GTCTCATGCGCTCAACCGTGGCCCAATAAACAGCATGCGCCAGATTAAGAGGATATCAGGTG 2400
QY 2401 CATATATGTAGGATGGAATAATGGTGAAGCCCAAGGSCCAGCCCATCTGAAAGATGTCATG 2460
DB 2401 CATATATGTAGGATGGAATAATGGTGAAGCCCAAGGSCCAGCCCATCTGAAAGATGTCATG 2460
QY 2461 CTGGCTGATGCAAGTGGGAATTTGATGATACTACTGAACATGATCATTTCTGGG 2520
DB 2461 CTGGCTGATGCAAGTGGGAATTTGATGATACTACTGAACATGATCATTTCTGGG 2520
QY 2521 CTCAGCCTGAAACTTCTCTACTCCCTCAACCGTCAAGCCTTACACAAACAAAGGAGATGGT 2580
DB 2521 CTCAGCCTGAAACTTCTCTACTCCCTCAACCGTCAAGCCTTACACAAACAAAGGAGATGGT 2580
QY 2581 GCTGCACAGCCCAACTGGTGTCCACCACTGGGGCAGTTCAGGGGAACCTCGGCTT 2640
DB 2581 GCTGCACAGCCCAACTGGTGTCCACCACTGGGGCAGTTCAGGGGAACCTCGGCTT 2640
QY 2641 GTGATTAACCACTCAGATGAATACTGCTCTTATTTAGTGGCACCCTCCGCTGGACACA 2700
DB 2641 GTGATTAACCACTCAGATGAATACTGCTCTTATTTAGTGGCACCCTCCGCTGGACACA 2700
QY 2701 TTTGACCTCTTCAGGGCTACCGTCTAAATTTGGCCGCAAGGATATGGAGCCACTTACT 2760
DB 2701 TTTGACCTCTTCAGGGCTACCGTCTAAATTTGGCCGCAAGGATATGGAGCCACTTACT 2760
QY 2761 ACTCTTGAGTCTCTGAAAGAGATCACTTTTACAGCTACAGCATCCACAAAGGAGCA 2820
DB 2761 ACTCTTGAGTCTCTGAAAGAGATCACTTTTACAGCTACAGCATCCACAAAGGAGCA 2820
QY 2821 TCATACGCTCTTCAGGCTCTCAGCCAGAAACAAAGTGGGCTTTGGGGAGGAGATGGTGAAG 2880
DB 2821 TCATACGCTCTTCAGGCTCTCAGCCAGAAACAAAGTGGGCTTTGGGGAGGAGATGGTGAAG 2880
QY 2881 GAGATTTCCATTCAGAGAAAGTACCAACTGGATTCCTCTCAAAACCTTTCACTCAGAGGC 2940
DB 2881 GAGATTTCCATTCAGAGAAAGTACCAACTGGATTCCTCTCAAAACCTTTCACTCAGAGGC 2940
QY 2941 ACCACTTCAACCTCGCTCCAGTTACTTTGGCAACCACTGTCCTGGCAGAGGAAATGGC 3000
DB 2941 ACCACTTCAACCTCGCTCCAGTTACTTTGGCAACCACTGTCCTGGCAGAGGAAATGGC 3000
QY 3001 ATTATCAACCAAGTATACCTTTCTTTATAGGATATCAACATCCCTCTCTCCCGATGGAG 3060
DB 3001 ATTATCAACCAAGTATACCTTTCTTTATAGGATATCAACATCCCTCTCTCCCGATGGAG 3060
QY 3061 CAGCTTATGTTCCAGCTGACACCACTATGACACTCTGAGGATATCAACATCCCTCTCTCCCGATGGAG 3120
DB 3061 CAGCTTATGTTCCAGCTGACACCACTATGACACTCTGAGGATATCAACATCCCTCTCTCCCGATGGAG 3120
QY 3121 TACGATTAAGATACGTCATACGAGCAAGAGGCCCCGGGCCATATAGTCCCAAGTGC 3180
DB 3121 TACGATTAAGATACGTCATACGAGCAAGAGGCCCCGGGCCATATAGTCCCAAGTGC 3180
QY 3181 CAGTTTCAGGACACTGCTGTGGATCAAGTGTTCGAAAAATTTTTCATGTCAAAAGCAGTA 3240
DB 3181 CAGTTTCAGGACACTGCTGTGGATCAAGTGTTCGAAAAATTTTTCATGTCAAAAGCAGTA 3240

QY 3241 ATGAAGACTTCCGTGTTGCTCTCTTGGGAGATTCAGAGAAATTAATACTCCGCCAATGCT 3300
DB 3241 ATGAAGACTTCCGTGTTGCTCTCTTGGGAGATTCAGAGAAATTAATACTCCGCCAATGCT 3300
QY 3301 TTCAAAATTTCTTTATGATGATGGGAAATGCTAGAAAGTGGATGGCCGAGCCACACAG 3360
DB 3301 TTCAAAATTTCTTTATGATGATGGGAAATGCTAGAAAGTGGATGGCCGAGCCACACAG 3360
QY 3361 AAGTTAAATTTCAACCTGGAACCTGAGAAATCATATTTCTTTGTCTGACAAATCGTGA 3420
DB 3361 AAGTTAAATTTCAACCTGGAACCTGAGAAATCATATTTCTTTGTCTGACAAATCGTGA 3420
QY 3421 AACAGTGTGTGGCTGTCAGCACAGGCTCAAGGAAAGACTGACACAGATGTTACGT 3480
DB 3421 AACAGTGTGTGGCTGTCAGCACAGGCTCAAGGAAAGACTGACACAGATGTTACGT 3480
QY 3481 ACCAAGCTGCTCTTCAATTTGGGAAGCAACCTTGGATGSCATGTTACTTGTGCACTGCT 3540
DB 3481 ACCAAGCTGCTCTTCAATTTGGGAAGCAACCTTGGATGSCATGTTACTTGTGCACTGCT 3540
QY 3541 GAAGTACTCTGCAATGAGAAATATAAAGGTTACTACATAAATTTGTGCTCTTTGAAGAA 3600
DB 3541 GAAGTACTCTGCAATGAGAAATATAAAGGTTACTACATAAATTTGTGCTCTTTGAAGAA 3600
QY 3601 TCTCGCGGAAATTTTATCAAGCCATGGAGAGTCCAGATGAAATGGAATAGATGAGCTG 3660
DB 3601 TCTCGCGGAAATTTTATCAAGCCATGGAGAGTCCAGATGAAATGGAATAGATGAGCTG 3660
QY 3661 CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGGAGAGATTTGAATTAAG 3720
DB 3661 CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGGAGAGATTTGAATTAAG 3720
QY 3721 CCATATATTTGCGCTCACTTTGATGTCCTTCCCACTGAGTTTCACTGGGGGATGACAAAG 3780
DB 3721 CCATATATTTGCGCTCACTTTGATGTCCTTCCCACTGAGTTTCACTGGGGGATGACAAAG 3780
QY 3781 CATTTATGGTGAATTTACAAACAGCAACTCTCAAAAGTGGTCAAGAAATATGTTCTTTGTG 3840
DB 3781 CATTTATGGTGAATTTACAAACAGCAACTCTCAAAAGTGGTCAAGAAATATGTTCTTTGTG 3840
QY 3841 TTACAGTAAATGGAACATGACAGAGTCTAAGATGATGCAACAGCCCTTACTCCGACCCC 3900
DB 3841 TTACAGTAAATGGAACATGACAGAGTCTAAGATGATGCAACAGCCCTTACTCCGACCCC 3900
QY 3901 GTGTGTCAATGGATCTGGATCCGACCAATCAAGTGAAGAGAGGCTTGTATCTGG 3960
DB 3901 GTGTGTCAATGGATCTGGATCCGACCAATCAAGTGAAGAGAGGCTTGTATCTGG 3960
QY 3961 GTTGTAGGTCTGCTTGTGAGTGTCTTTATCATCTGCAATTTGCTTCTTCTTCTT 4020
DB 3961 GTTGTAGGTCTGCTTGTGAGTGTCTTTATCATCTGCAATTTGCTTCTTCTTCTT 4020
QY 4021 TATAAAGGAGAGGCGAGAGTCCGACTCTAGAAAAAGCAGCATACCGAACAAATAGGAG 4080
DB 4021 TATAAAGGAGAGGCGAGAGTCCGACTCTAGAAAAAGCAGCATACCGAACAAATAGGAG 4080
QY 4081 ATCCCTTTCACACCAACCAAGAGCCCTGTAGAACTGAGGCGCTTAACTTTCAAAACCG 4140
DB 4081 ATCCCTTTCACACCAACCAAGAGCCCTGTAGAACTGAGGCGCTTAACTTTCAAAACCG 4140
QY 4141 GGTATGCTAGCCATCTCTCAATACCCATCTTGGAACTTGGACCAACATTTGAAGATTG 4200
DB 4141 GGTATGCTAGCCATCTCTCAATACCCATCTTGGAACTTGGAACTTGGAAAGATTG 4200
QY 4201 AAAGCAATGCAACTTTGAAGTTTCCAGGAATATGAGTCAATTTGACCTTGGCCAGCAG 4260
DB 4201 AAAGCAATGCAACTTTGAAGTTTCCAGGAATATGAGTCAATTTGACCTTGGCCAGCAG 4260
QY 4261 TTCACTTTGGGAACATTTCAAACTTTGGAGTAAACAAACCAAGAAATAGATACGCGAATGTA 4320
DB 4261 TTCACTTTGGGAACATTTCAAACTTTGGAGTAAACAAACCAAGAAATAGATACGCGAATGTA 4320
QY 4321 ATGCGATATGATCATTTCCCGGGTTCTCTATACAGCTATAGAGGGATTCACGAGGATGAC 4380

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2005, 22:42:42 ; Search time 3361 Seconds
(without alignments)
11687.798 Million cell updates/sec

Title: US-10-772-636-63
Perfect score: 6263
Sequence: 1 gctaactcaagggagcgtc.....aagtgaatttttaaaagttt 6263

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	6263	100.0	6263 18	US-10-641-643-1473
2	6263	100.0	6263 21	Sequence 1473, Ap
3	2473	39.5	6363 18	US-10-772-636-63
4	2359	37.7	7741 17	US-10-641-643-879
5	2357.4	37.6	7702 11	US-10-291-265-14
6	2357.4	37.6	7702 17	US-09-719-272-3
7	2357.4	37.6	7702 18	US-10-210-838-4
				Sequence 1336, Ap

8	2357.4	37.6	7702	21	US-10-945-647-3	Sequence 3, Appli
9	2357.4	37.6	7702	21	US-10-482-029-51	Sequence 51, Appli
10	2357.4	37.6	7702	21	US-10-945-772-3	Sequence 3, Appli
11	2357.4	37.6	7702	21	US-10-505-680-391	Sequence 391, App
12	2357.4	37.6	7718	14	US-10-175-523-192	Sequence 192, App
13	2357.4	37.6	7718	17	US-10-133-937-9	Sequence 9, Appli
14	2357.4	37.6	7718	17	US-10-159-563-9	Sequence 9, Appli
15	2357.4	37.6	7724	17	US-10-210-838-18	Sequence 18, Appli
16	2357.4	37.6	7945	20	US-10-357-930-27910	Sequence 27910, A
17	2348	37.5	7705	17	US-10-291-265-486	Sequence 486, App
18	2307.2	36.8	6734	21	US-10-764-420-2695	Sequence 2695, Ap
19	2303.2	36.8	8232	14	US-10-198-846-10976	Sequence 10976, A
20	1345.8	21.5	4078	17	US-10-305-720-1132	Sequence 1132, Ap
21	1290	20.6	6143	17	US-10-210-838-11	Sequence 11, Appli
22	1245.2	19.9	5412	17	US-10-258-666-11	Sequence 11, Appli
23	1196.6	19.1	4190	17	US-10-388-934-253	Sequence 253, App
24	1095.8	17.5	3467	11	US-09-719-272-1	Sequence 1, Appli
25	1095.8	17.5	3467	21	US-10-945-647-1	Sequence 1, Appli
26	1095.8	17.5	3467	21	US-10-945-772-1	Sequence 1, Appli
27	829.4	13.2	3064	9	US-09-925-300-10	Sequence 10, Appli
28	442	7.1	2872	21	US-10-777-144-4	Sequence 4, Appli
29	442	7.1	2872	21	US-10-777-145-4	Sequence 4, Appli
30	442	7.1	2872	21	US-10-777-186-4	Sequence 4, Appli
31	438.8	7.0	2409	21	US-10-777-144-2	Sequence 2, Appli
32	438.8	7.0	2409	21	US-10-777-145-2	Sequence 2, Appli
33	438.8	7.0	2430	17	US-10-777-186-2	Sequence 2, Appli
34	438.8	7.0	2430	17	US-10-210-556-227	Sequence 227, App
35	438.8	7.0	2526	21	US-10-956-157-4701	Sequence 4701, Ap
36	438.8	7.0	2565	17	US-10-305-720-1311	Sequence 1311, Ap
37	438.8	7.0	2565	17	US-10-210-556-21	Sequence 21, Appli
38	438.8	7.0	2826	17	US-10-264-049-963	Sequence 963, App
39	438.8	7.0	3148	17	US-10-210-556-4	Sequence 4, Appli
40	438.8	7.0	3279	17	US-10-210-556-226	Sequence 226, App
41	438.8	7.0	3615	17	US-10-305-720-1254	Sequence 1254, Ap
42	438.8	7.0	3615	17	US-10-210-556-18	Sequence 18, Appli
43	438.8	7.0	3643	17	US-10-210-556-225	Sequence 225, App
44	438.8	7.0	3643	21	US-10-956-157-947	Sequence 947, App
45	438.8	7.0	3643	21	US-10-956-157-948	Sequence 948, App

ALIGNMENTS

RESULT 1

US-10-641-643-1473
; Sequence 1473, Application US/10641643
; Publication No. US20040077003A1

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.

Susan G. Stuart

Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL

GENE EXPRESSION

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/641.643

FILING DATE: 14-Aug-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>


```
;
;
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
;
; INFORMATION FOR SEQ ID NO: 1473:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 6263 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: 9755652
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1473 :
US-10-641-643-1473

Query Match      100.08; Score 6263; DB 18; Length 6263;
Best Local Similarity 100.08; Pred. No. 0;
Matches 6263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTAACTCAAGGAGACGCTCTGGTGAACACCCGCTGGGATCTAAAGAACAAAGCTCTGAAAG 60
DB 1 GCTAACTCAAGGAGACGCTCTGGTGAACACCCGCTGGGATCTAAAGAACAAAGCTCTGAAAG 60
QY 61 TGTTCAGCTGAAATTTTCAGATCGGACAGACTCGCTCGCGTCCGGAGGACGATGATTTCCA 120
DB 61 TGTTCAGCTGAAATTTTCAGATCGGACAGACTCGCTCGCGTCCGGAGGACGATGATTTCCA 120
QY 121 AGCTGCTCGGCACGCTGCTGCCAAGCTGACAGATGGTGCAGCTAGCAGGCTCTGCTG 180
DB 121 AGCTGCTCGGCACGCTGCTGCCAAGCTGACAGATGGTGCAGCTAGCAGGCTCTGCTG 180
QY 181 CTGCTCTCTCACTTTCTCTCGCACGATGCTGAGACACCTCCAAAGGTTTACACGAACA 240
DB 181 CTGCTCTCTCACTTTCTCTCGCACGATGCTGAGACACCTCCAAAGGTTTACACGAACA 240
QY 241 CCGGTGTGATCAGACAGGGGTCTTGGCGGAGTGTGCTCTTTTCATCTGCGCAAGCTACGGGA 300
DB 241 CCGGTGTGATCAGACAGGGGTCTTGGCGGAGTGTGCTCTTTTCATCTGCGCAAGCTACGGGA 300
QY 301 GACCCACAGACCTAAATTTGCTGGACAAAGGAGGAGGAGTGCAGCAATCAGAGATTT 360
DB 301 GACCCACAGACCTAAATTTGCTGGACAAAGGAGGAGGAGTGCAGCAATCAGAGATTT 360
QY 361 GAGGTAAATAGAGTTTGACGATGGGTCTGGATCAGTTCTCAGAAATACACCCCTTACGGACT 420
DB 361 GAGGTAAATAGAGTTTGACGATGGGTCTGGATCAGTTCTCAGAAATACACCCCTTACGGACT 420
QY 421 CCGAGGGATAGGCCATTTATGAATGTGTGGCTCAAAATATGTGGGAGAAATTAAGTGA 480
DB 421 CCGAGGGATAGGCCATTTATGAATGTGTGGCTCAAAATATGTGGGAGAAATTAAGTGA 480
QY 481 TCCACACAGCTCAGAGTTTTCGGGAGATCAAAATCCAGGGGCTTCCCTACCATTTGAC 540
DB 481 TCCACACAGCTCAGAGTTTTCGGGAGATCAAAATCCAGGGGCTTCCCTACCATTTGAC 540
QY 541 ATGGGCCCCACAGTTGAAGGTGTTGAGCGTACTCGCAGGGCCACCATGCTTTGTGCAGCC 600
DB 541 ATGGGCCCCACAGTTGAAGGTGTTGAGCGTACTCGCAGGGCCACCATGCTTTGTGCAGCC 600
QY 601 AGTGTAAATCCGGATCCAGAAATCACTGGTTTAAAGATTTTCTACCTGTGGACACAGC 660
DB 601 AGTGTAAATCCGGATCCAGAAATCACTGGTTTAAAGATTTTCTACCTGTGGACACAGC 660
QY 661 AACACAAATGTCGATTAAGCAGTTACGATCAGAAATCTATTGGTGTGACACCAATAGA 720
DB 661 AACACAAATGTCGATTAAGCAGTTACGATCAGAAATCTATTGGTGTGACACCAATAGA 720
QY 721 GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGAGTGTGTGCC 780
DB 721 GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGAGTGTGTGCC 780
781 ACCAACAGCGCGGCACTCGCTATTTCGGCTCTCGCCTGCTGCAATTTATATGTCAGAGAGCTCGGA 840
781 ACCAACAGCGCGGCACTCGCTATTTCGGCTCTCGCCTGCTGCAATTTATATGTCAGAGAGCTCGGA 840
841 GAAGTTTCGGCGTCTCCACCAAGATTTCTATCCACCCCACTAATCATGAAATCATGCCA 900
841 GAAGTTTCGGCGTCTCCACCAAGATTTCTATCCACCCCACTAATCATGAAATCATGCCA 900
901 GCGGGAAGCGTTAATATCACCTGTGTGGCGTGGGTCACCAATGCTTATGTAAGTGG 960
901 GCGGGAAGCGTTAATATCACCTGTGTGGCGTGGGTCACCAATGCTTATGTAAGTGG 960
961 ATGTTGGGGGACAGAGATCTGACACCTGGAAGATGATATGCCAATAGGAAGAAATGTCCTA 1020
961 ATGTTGGGGGACAGAGATCTGACACCTGGAAGATGATATGCCAATAGGAAGAAATGTCCTA 1020
1021 GAACTGAATGATGTAAGACAGTCAAGCAATTTACACCTGTGTGCTATGTCAACACTGGGT 1080
1021 GAACTGAATGATGTAAGACAGTCAAGCAATTTACACCTGTGTGCTATGTCAACACTGGGT 1080
1081 GTCAATTGAAGCAATAGCACAGATCACTGTCAAAGCCTTACCCAAACCTCCAGGAACCTCT 1140
1081 GTCAATTGAAGCAATAGCACAGATCACTGTCAAAGCCTTACCCAAACCTCCAGGAACCTCT 1140
1141 GTAGTGACCGAGACGAGCTACAGCATCAAGCATCACACTGACGTGGGACTCTGGGAAACCTTGAG 1200
1141 GTAGTGACCGAGACGAGCTACAGCATCAAGCATCACACTGACGTGGGACTCTGGGAAACCTTGAG 1200
1201 CCTGTTTCTTATTACATAAATTCAGCATATAAATCTAAAGAACTCTGAGGAACCTTTACAAAGAA 1260
1201 CCTGTTTCTTATTACATAAATTCAGCATATAAATCTAAAGAACTCTGAGGAACCTTTACAAAGAA 1260
1261 ATTGATGGGTGGCGACCAACACGCTACAGTGTGCTGGACTAAGTCTCCTACTCGGATAT 1320
1261 ATTGATGGGTGGCGACCAACACGCTACAGTGTGCTGGACTAAGTCTCCTACTCGGATAT 1320
1321 GAATTCAGGGTGTGTGCTGCTCAATTAATTTGGGGGGGGGCTCCACAGGAACTGTGCTA 1380
1321 GAATTCAGGGTGTGTGCTGCTCAATTAATTTGGGGGGGGGCTCCACAGGAACTGTGCTA 1380
1381 ACACAAACCTCAGAGCAAGCAACCATCCAGTGTCCCGCAGGGGATGTCCAGGCAACGAATGTTG 1440
1381 ACACAAACCTCAGAGCAAGCAACCATCCAGTGTCCCGCAGGGGATGTCCAGGCAACGAATGTTG 1440
1441 AGTTTCGACCAACATTTTGGTGTAGTGAAGAAACCTGAAAGGCCAAATGACAGATCAA 1500
1441 AGTTTCGACCAACATTTTGGTGTAGTGAAGAAACCTGAAAGGCCAAATGACAGATCAA 1500
1501 GGATATAGAGTTTATATACAATGGATCCCACTCAACATGTCAACAACTGGATGAAACAC 1560
1501 GGATATAGAGTTTATATACAATGGATCCCACTCAACATGTCAACAACTGGATGAAACAC 1560
1561 AATGTAGCTGACAGCCAAATCACTACTATTATTTGGCAACTTAGTGCCCGCAGAAACATATTTCT 1620
1561 AATGTAGCTGACAGCCAAATCACTACTATTATTTGGCAACTTAGTGCCCGCAGAAACATATTTCT 1620
1621 GTCAAGTCTCTGGCTTTTACCTCAATTTGGAGATGGTCCCTTTCAAGTGACATACAAGTC 1680
1621 GTCAAGTCTCTGGCTTTTACCTCAATTTGGAGATGGTCCCTTTCAAGTGACATACAAGTC 1680
1681 ATCACTCAGACAGAGTACAGGGGACGACTAACTTCAAGCAGAACTCTGAGTCTGAA 1740
1681 ATCACTCAGACAGAGTACAGGGGACGACTAACTTCAAGCAGAACTCTGAGTCTGAA 1740
1741 ACAAGTATTTTGTCTCTCTTGGACACCTCCAGTTTCAGATACCATTTGCAACTATGAATG 1800
1741 ACAAGTATTTTGTCTCTCTTGGACACCTCCAGTTTCAGATACCATTTGCAACTATGAATG 1800
1801 GTCTACAAAGATGGGAGCATGGAGAGGAGCAACGAATTTACCATTTGAGCCAGGACATCA 1860
1801 GTCTACAAAGATGGGAGCATGGAGAGGAGCAACGAATTTACCATTTGAGCCAGGACATCA 1860
```

QY	1861	TATAGGCTGCAAGGACTGAAACCAACAGCTTATATCTATTTCCGCTGCTGCTGCAAGCTCC	1920
Db	1861		
QY	1921	CCTCAAGGCTGGGTCTTCTACTCGAGAAATATCAGCTAGAACCATCGAGTCAAGCGG	1980
Db	1921		
QY	1981	TCAGCTCTCTCTCAAGACATTTAGTTGCAACAGCCCAAGTTCCACTAGTATTTTGTAAAGT	2040
Db	1981		
QY	2041	TGGCAACCTCTCAAGAGCATTTAGTTGCAACAGCCCAAGTTCCACTAGTATTTTGTAAAGT	2100
Db	2041		
QY	2101	ACTGCAGTGGATGGGAGATGACAAAGCTCACAGAGATTTTGGGAATTTCCCTTCGGACACT	2160
Db	2101		
QY	2161	ACCAATACCTTTTGGAAACAGCTGGAAATATGCACTGGAATACCGGATCACTGTGACAGCC	2220
Db	2161		
QY	2221	CATACAGATGTGGCCCTGGCCCTGAGAGCTTTGTCGGTGTGATTTGGAACCAATGAAGAT	2280
Db	2221		
QY	2281	GTTCCCTAGTGTCTCTCGCAAGCTGAGAGCTTTGTCGGTGTGATTTGGAACCAATGAAGAT	2340
Db	2281		
QY	2341	GTCTCATGGCGCTCACCGCTGCCAATAAAGACATGAGAGCTGCACTCAACATCTGTTAAA	2400
Db	2341		
QY	2401	CATTATGTAGAGATGGAAATTTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG	2460
Db	2401		
QY	2461	CTGGCTGATGACAGTGGAAATTTGATGATCTACTGAAACATGACATCATTTCTGGG	2520
Db	2461		
QY	2521	CTCCAGCTGAACTTCTACTCTCTCAAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG	2580
Db	2521		
QY	2581	GCTGCAGAGCCCAAACTGGTGTCCACCACTGGGGCAGTTTCCAGGGAAACCTCGGCTT	2640
Db	2581		
QY	2641	GTGATTAACCACTCAGATGAATCTCTTATTTAGTGGCACTTCCAGGGAAACCTCGGCTT	2700
Db	2641		
QY	2701	TTTGGACCTCTTCAGGGCTACCGTCTAAATTTTGGCCGCAAGGATATGGAGCCACTTACT	2760
Db	2701		
QY	2761	ACTCTTGAGTCTCTGAAAGAGAGATCACTTTTACAGCTACAGATCAAGAGGAGCA	2820
Db	2761		
QY	2821	TCATACGCTCTCAGGCTCTCAGCAGAAACAAAGTGGGCTTTGGGGAGGATGGTGAAG	2880
Db	2821		
QY	2881	GAGATTTCCATTTCCAGAGAGAGTACCACTGGATTTCCCTCAAAACCTTCACTCAGAGGC	2940
Db	2881		

QY	2941	ACCACTTCAACCTCCGCTCAGTTATCTTTGGCAACCACTGCTGCTGGCAGAGAGAAATGGC	3000
Db	2941		
QY	3001	ATTATCAACCAAGTATACCTTCTTTATAGGATATCAACATCCCTCTTCCCGATGAG	3060
Db	3001		
QY	3061	CAGCTTATTTGTTCCAGCTGACCACTATGACCTCACTGCTTTAAACCCAGATACCAACA	3120
Db	3061		
QY	3121	TACGATGTAAAGTACGCTCATACGAGCAAGGGCCCGGCCATATAGTCCCGATGTC	3180
Db	3121		
QY	3181	CAGTTTCAGGACACCTGCTGATCAAGTGTTCGAAAAATTTTCATGTCAAAAGCAGTA	3240
Db	3181		
QY	3241	ATGAAGACTTCCGCTGCTGCTTTGGGAGATTTCCAGAGAAATTAATCTCCGCCATGCT	3300
Db	3241		
QY	3301	TTCAAAATTTCTTTATGATGATGGGAAATGTTAGAGAGTGGATGGCCGAGCCACACAG	3360
Db	3301		
QY	3361	AAGTTAATTTGTCACCTGMAAGCTGAGAAATCATATTCTTTGCTGACAAAATCGTGA	3420
Db	3361		
QY	3421	AACAGTCTGCTGGGCTGCAGCACAGGGTCAACGGCAAGAGCTGCACCAAGATGTATTACGT	3480
Db	3421		
QY	3481	ACCAAGCTTGCCTTCATTGGGAGACCACTTTGATGGCATGATTTACTGTGCAATGCT	3540
Db	3481		
QY	3541	GAAGTACCTGCAATGAGAAATATAAAGGTTTACTACATAATATTTGTCCTTTGAAGAAA	3600
Db	3541		
QY	3601	TCTCGCGGAAATTTTATCAAGCCATGGGAGAGTCCAGATGAAATGGAATTAGATGAGCTG	3660
Db	3601		
QY	3661	CTTAAGGAGATATCTAGAGAGCCGAGAGCAATCCGTTATGGGAGAGAGTTGAATTAAG	3720
Db	3661		
QY	3721	CCATATATTCGCGCTCACTTTGATGCTCTTCCCACTGAGTTTCACTGGGGGATGACAAAG	3780
Db	3721		
QY	3781	CATTATGCTGGATTTTCAAAACAGCAACTCCAAAGTGGTCAAGAATATGCTCTCTTTG	3840
Db	3781		
QY	3841	TTAGCAGTATGGAACATGAGAGTCTAAGATGATGCAACAGCCCTTACTCCGACCCC	3900
Db	3841		
QY	3901	GTGCTGTCAATGGATCTGGATCCGAGCCCAATCAGGATGAAGAAGAGCTTCATCTGG	3960
Db	3901		
QY	3961	GTTGTAGGCTCTGCTTTCAGTGGTCTTTTATCATCTGCAATTTGCTATTGCTTTCTTT	4020
Db	3961		
QY	4021	TATAAAGGAGAGGGCAGAGTCCGACTCTAGAAAAAGCAGCATACCGCAACAATAAGAG	4080

Db 4021 TATAAAGAGAGGGGAGAGTCCGACTCTAGAAAAAGCAGCATACCGAACAATAGAGAG 4080
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4081 ATCCCTTTACACACCAACCAACAGACCCCTGTGAACTGAGGGCCCTTAACCTTTCAAAACCG 4140
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4081 ATCCCTTTACACACCAACCAACAGACCCCTGTGAACTGAGGGCCCTTAACCTTTCAAAACCG 4140
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4141 GGTATGGTGTAGCCATCTCCATACCCATCTTGAACTTGGAACCTTGAGACCAATTTGAAAGATTG 4200
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4141 GGTATGGTGTAGCCATCTCCATACCCATCTTGAACTTGGAACCTTGAGACCAATTTGAAAGATTG 4200
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4201 AAAGCAATGACAACTTTGAAGTTTCCAGGAATATGAGTCAATTTGACCTTGGCCAGCAG 4260
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4201 AAAGCAATGACAACTTTGAAGTTTCCAGGAATATGAGTCAATTTGACCTTGGCCAGCAG 4260
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4261 TTCACTTTGGGAACATTCANAACCTTGGAAAGTAAACAAACCAAGAAATAGATACGCGAATGTA 4320
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4261 TTCACTTTGGGAACATTCANAACCTTGGAAAGTAAACAAACCAAGAAATAGATACGCGAATGTA 4320
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4321 ATCCCATATGATCAATTCCTGGGTTCTCCTATCAGCTATAGAGGGATCCCAAGGATGAC 4380
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4321 ATCCCATATGATCAATTCCTGGGTTCTCCTATCAGCTATAGAGGGATCCCAAGGATGAC 4380
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4381 TATGTGAATGCCAATCATAGATGGGTATAGGAAGCAAAATGCTATATTTGCAACACAG 4440
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4381 TATGTGAATGCCAATCATAGATGGGTATAGGAAGCAAAATGCTATATTTGCAACACAG 4440
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4441 GGATCTCTCCCGGAAACATTTGGGGACTTTTGGGAGATGATATGGAACCAACCGAGTGCC 4500
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4441 GGATCTCTCCCGGAAACATTTGGGGACTTTTGGGAGATGATATGGAACCAACCGAGTGCC 4500
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4501 ACAGTTGTATGATGACAAAACCTAGAGAAAGATCAAGGGTGAAGTGTGACCAAGTATGG 4560
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4501 ACAGTTGTATGATGACAAAACCTAGAGAAAGATCAAGGGTGAAGTGTGACCAAGTATGG 4560
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4561 CCTAGCAGAGCAGAGAAACCCAGGACTCGTTCAGAGTAAAGCTGTGATCTGTGGAG 4620
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4561 CCTAGCAGAGCAGAGAAACCCAGGACTCGTTCAGAGTAAAGCTGTGATCTGTGGAG 4620
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4621 CTGGCCACATATGTGTTGGAACATTTGCACTTTTACAAAGATGTTCAAGTGAAGAGAGA 4680
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4681 GAAGTGAGACAAATTCAGTTTCAACGCTGGGCTGATCATGTGTTTCCAGACACCCCTACA 4740
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4681 GAAGTGAGACAAATTCAGTTTCAACGCTGGGCTGATCATGTGTTTCCAGACACCCCTACA 4740
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4741 CCTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATG 4800
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4741 CCTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATG 4800
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4801 GTTGTGCACTGCAGTGGGGAGTTGGCCGAGTGGTTCCTTCATCGTCATAGATGCCATG 4860
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4801 GTTGTGCACTGCAGTGGGGAGTTGGCCGAGTGGTTCCTTCATCGTCATAGATGCCATG 4860
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4861 TTAGAAAGAAATAGCATGAAACCTGTAGATATTTATGGCCATGTAACCTTTAATGAGA 4920
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4861 TTAGAAAGAAATAGCATGAAACCTGTAGATATTTATGGCCATGTAACCTTTAATGAGA 4920
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4921 GCCCAGAGAACTATATGGTTTCAACAGAGAACCAATACATCTTTATCCATGATGCACTG 4980
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4921 GCCCAGAGAACTATATGGTTTCAACAGAGAACCAATACATCTTTATCCATGATGCACTG 4980
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4981 TTAGAGCAGTGTGTTGGAAATACCGAAGTGGCAGCTAGAAAACCTGTATGCTCATACT 5040
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4981 TTAGAGCAGTGTGTTGGAAATACCGAAGTGGCAGCTAGAAAACCTGTATGCTCATACT 5040
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5041 CAGAAGCTGACAAAATAGAACGGGAGAGATGTCACAGAAATGGAGCTCGAATTTAAG 5100
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5041 CAGAAGCTGACAAAATAGAACGGGAGAGATGTCACAGAAATGGAGCTCGAATTTAAG 5100
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5101 CGTCTAGCCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGGCCAAATCTTCCATGTAAT 5160
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 5101 CGTCTAGCCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCCAAATCTTCCATGTAAT 5160
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5161 AAATTCAAAATCGCCTTGTTAATATATGCCATATGAATCCCAAGGGTATGCTCTCAG 5220
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5161 AAATTCAAAATCGCCTTGTTAATATATGCCATATGAATCCCAAGGGTATGCTCTCAG 5220
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5221 CCTATCCGTGGAGTAGAAGGATCTGATTACATCAATGCCAGTTTATTGATGATACAGA 5280
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5221 CCTATCCGTGGAGTAGAAGGATCTGATTACATCAATGCCAGTTTATTGATGATACAGA 5280
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5281 CAAAGAGAGCCTACATCGCTACCCAGGGGCCCTTGGCAGAGACCACTGAAAGCTTCTGG 5340
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5281 CAAAGAGAGCCTACATCGCTACCCAGGGGCCCTTGGCAGAGACCACTGAAAGCTTCTGG 5340
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5341 CGATGCTCTGGGAAACAAATTCACCATATGTTGTGATGCTCACCAGCTGCGTGAATG 5400
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5341 CGATGCTCTGGGAAACAAATTCACCATATGTTGTGATGCTCACCAGCTGCGTGAATG 5400
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5401 GGCAGAGAGAAATGTCAACCAATCTGGCCAGAGAACGGTCTGCAAGATACCACTACTTT 5460
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5401 GGCAGAGAGAAATGTCAACCAATCTGGCCAGAGAACGGTCTGCAAGATACCACTACTTT 5460
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5461 GTTGTAGATCCCATGGCTGAGTACAAATGCCACAGTATATCTTAAGGGAATCAAGGTC 5520
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5461 GTTGTAGATCCCATGGCTGAGTACAAATGCCACAGTATATCTTAAGGGAATCAAGGTC 5520
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5521 ACAGATGCCAGGAGCGCCAGTCCGNAACAGTAAAGCAGTTCAGTTCACTGACTGCGCA 5580
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5521 ACAGATGCCAGGAGCGCCAGTCCGNAACAGTAAAGCAGTTCAGTTCACTGACTGCGCA 5580
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5581 GAGCAAGGAGTGCACAAAGTCCGAGAGAGGATTTATTGACTTCATCGGCCAAGTCCATAAA 5640
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5581 GAGCAAGGAGTGCACAAAGTCCGAGAGAGGATTTATTGACTTCATCGGCCAAGTCCATAAA 5640
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5641 AAAAAAGAACAGTTTGGCCAAAGTGAACCCATTTTCACTGCTCCATTCGACGCGGGCTTGA 5700
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5641 AAAAAAGAACAGTTTGGCCAAAGTGAACCCATTTTCACTGCTCCATTCGACGCGGGCTTGA 5700
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5701 AGAATCTGGAGTCTTCATACGCTTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGGATT 5760
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5701 AGAATCTGGAGTCTTCATACGCTTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGGATT 5760
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5761 GTAGATATCTTCCAGACTGTCAAAATGTTAAGAACCAACAGCAGCAGTATGGTACAGACA 5820
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5761 GTAGATATCTTCCAGACTGTCAAAATGTTAAGAACCAACAGCAGCAGTATGGTACAGACA 5820
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5821 GAGGATCAATATCAGTTTCTCTATCGTCCGCACTAGAGTACCTGGGCGAGCTTTGACAC 5880
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5821 GAGGATCAATATCAGTTTCTCTATCGTCCGCACTAGAGTACCTGGGCGAGCTTTGACAC 5880
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5881 TATGCAACGTAGAAACCCCTGACCCATCTCTGGATTTTATACAGGCCCTTCAATATCCA 5940
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5881 TATGCAACGTAGAAACCCCTGACCCATCTCTGGATTTTATACAGGCCCTTCAATATCCA 5940
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5941 TGGAGTCTCTTCTGAGCCATACAGGGCACTTTGAGAAGTCTCTTTAACTTCTAGCTAAACA 6000
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5941 TGGAGTCTCTTCTGAGCCATACAGGGCACTTTGAGAAGTCTCTTTAACTTCTAGCTAAACA 6000
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
6001 ACTACTTGTAGGGGACTATTACACAAAACAAATTTAAAAACAAATTTTTCAGGTGGACC 6060
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
6001 ACTACTTGTAGGGGACTATTACACAAAACAAATTTAAAAACAAATTTTTCAGGTGGACC 6060
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
6061 AAGAAATCTTTGACATCGCCCTTCCACCATCTGCTCATAATAAATTTTACGGGCA 6120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
6061 AAGAAATCTTTGACATCGCCCTTCCACCATCTGCTCATAATAAATTTTACGGGCA 6120
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
6121 AGGGGAGGGAATGTTTAAAAAGAAAGTCTTGAATTTAGTTTATTTAGTAAAGATAC 6180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
6121 AGGGGAGGGAATGTTTAAAAAGAAAGTCTTGAATTTAGTTTATTTAGTAAAGATAC 6180
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
6181 TGCTGACCTGTGCTTCAATTTCTAACTGTGTAACTTTTTTTTAAACAAAATGTATCTCG 6240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
6181 TGCTGACCTGTGCTTCAATTTCTAACTGTGTAACTTTTTTTTAAACAAAATGTATCTCG 6240

Qy 6241 ATAAAGTGAATTTTAAAAAGTT 6263
Db 6241 ATAAAGTGAATTTTAAAAAGTT 6263

RESULT 2

US-10-772-636-63

; Sequence 63, Application US/10772636
; Publication No. US20050042687A1
; GENERAL INFORMATION:
; APPLICANT: Kelly, Louise M.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Farlow, Deborah
; APPLICANT: Healy, Aileen
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 9118, 990, 17662, 81982, 630,
; TITLE OF INVENTION: 21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,
; TITLE OF INVENTION: 9465, 23544, 3666, 27417, 57259, 21844, 943, 2061, 5891,
; TITLE OF INVENTION: 9137, 13908, 14310, 17600, 25584, 27824, 28469, 38947,
; TITLE OF INVENTION: 53003, 965, 56639, 9661, 16052, 1521, 6662, 13913, 12405 OR
; TITLE OF INVENTION: 5014
; FILE REFERENCE: MPI03-01SP1RNOMIN
; CURRENT APPLICATION NUMBER: US/10/772,636
; CURRENT FILING DATE: 2004-02-05
; PRIOR APPLICATION NUMBER: US 60/445,241
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/448,389
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/456,320
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/460,279
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/465,924
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: US 60/470,052
; PRIOR FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US 60/498,106
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/500,179
; PRIOR FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: US 60/502,909
; PRIOR FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: US 60/510,351
; PRIOR FILING DATE: 2003-10-10
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 6263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (154)...(5892)
US-10-772-636-63

Query Match 100.0%; Score 6263; DB 21; Length 6263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTAACTCAAGGAGAGCTCTGGTGAACACCCGTGGGATCTAAAGAACAAAGCTCTGAAG 60
Db 1 GCTAACTCAAGGAGAGCTCTGGTGAACACCCGTGGGATCTAAAGAACAAAGCTCTGAAG 60
Qy 61 TGTTCAGCTCAATTTTCAGATCGGACAGACTCCGTGGGCTCCGGAGGAGATGATCCA 120
Db 61 TGTTCAGCTCAATTTTCAGATCGGACAGACTCCGTGGGCTCCGGAGGAGATGATCCA 120
Qy 121 AGCTGCTCGGCACGCTCTGCCAAGCTGACAGGATGGTGCAGTAGCCAGGCTCTGCTG 180
Db 121 AGCTGCTCGGCACGCTCTGCCAAGCTGACAGGATGGTGCAGTAGCCAGGCTCTGCTG 180

Qy 181 CTGCTCTCTCACTTTCTTCTCCGACGAGATGCTGAGACACCTCCAGGTTTACACGAACA 240
Db 181 CTGCTCTCTCACTTTCTTCTCCGACGAGATGCTGAGACACCTCCAGGTTTACACGAACA 240
Qy 241 CCGTTGATCAGACAGGGGTCTCTGGCGGAGTTCCTCTTTTCATCTGCAAGCTACGGGA 300
Db 241 CCGTTGATCAGACAGGGGTCTCTGGCGGAGTTCCTCTTTTCATCTGCAAGCTACGGGA 300
Qy 301 GACCAAGAGCCTAAATTTGTCTGGAACAAAAAGAAAGAAAGTTCAGCAATCAGAGATTT 360
Db 301 GACCAAGAGCCTAAATTTGTCTGGAACAAAAAGAAAGAAAGTTCAGCAATCAGAGATTT 360
Qy 361 GAGTAATAGAGTTTACAGATGGGTCTGGATCAGTCTCAGAAATCAACCTTACGGACT 420
Db 361 GAGTAATAGAGTTTACAGATGGGTCTGGATCAGTCTCAGAAATCAACCTTACGGACT 420
Qy 421 CCGAGGATGAGGCGCATTTTATGAATGTGCGCTCAATATATGTTGGGAGAAATTAAGTGA 480
Db 421 CCGAGGATGAGGCGCATTTTATGAATGTGCGCTCAATATATGTTGGGAGAAATTAAGTGA 480
Qy 481 TCCACCAAGCTCAGATTTTTCGCGGAGAGATCAATTTCCAGGGGCTTCCCTACCATTTGAC 540
Db 481 TCCACCAAGCTCAGATTTTTCGCGGAGAGATCAATTTCCAGGGGCTTCCCTACCATTTGAC 540
Qy 541 ATGGGCCCCACAGTTGAAGTGGTTGACGCTACTCGACGCGCCACCATGCTTTGTGACGCC 600
Db 541 ATGGGCCCCACAGTTGAAGTGGTTGACGCTACTCGACGCGCCACCATGCTTTGTGACGCC 600
Qy 601 AGTGGTAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC 660
Db 601 AGTGGTAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC 660
Qy 661 AACCAACAATCGTCTGATTAAGCAGTTTACGATCAGAAATCTATTGGTGGTACACCAATGA 720
Db 661 AACCAACAATCGTCTGATTAAGCAGTTTACGATCAGAAATCTATTGGTGGTACACCAATGA 720
Qy 721 GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGATGATGTTGCC 780
Db 721 GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGATGATGTTGCC 780
Qy 781 ACCAAGACGCGGGGACTCCGCTATTCCGCTCTCCCAATTTATATGTCAGAGAGCTCGGA 840
Db 781 ACCAAGACGCGGGGACTCCGCTATTCCGCTCTCCCAATTTATATGTCAGAGAGCTCGGA 840
Qy 841 GAAGTTCCGCTGTCCCAACCAAGATTTCTATCCCAACCACTTAATCATGAATCATGCCA 900
Db 841 GAAGTTCCGCTGTCCCAACCAAGATTTCTATCCCAACCACTTAATCATGAATCATGCCA 900
Qy 901 GCGGGAAGCGTTAATATACCTGTGCGCGGTGCGGTGTCACCAATGCTTATGTAAGTGG 960
Db 901 GCGGGAAGCGTTAATATACCTGTGCGCGGTGCGGTGTCACCAATGCTTATGTAAGTGG 960
Qy 961 ATGTTGGGGGACAGAGATCTGACACCTGAAGATGATATGCAATAGGAAGAAATGTCCTA 1020
Db 961 ATGTTGGGGGACAGAGATCTGACACCTGAAGATGATATGCAATAGGAAGAAATGTCCTA 1020
Qy 1021 GAAGTGAATGATGTAAGACAGTTCAGCAATTTACACCTGTGTGCTATGTCACCACTGGGT 1080
Db 1021 GAAGTGAATGATGTAAGACAGTTCAGCAATTTACACCTGTGTGCTATGTCACCACTGGGT 1080
Qy 1081 GTCAATGAAGCAATAGCAGATCAGTGTCAAGAGCTTACCACCACTCCAGGAGCTCCT 1140
Db 1081 GTCAATGAAGCAATAGCAGATCAGTGTCAAGAGCTTACCACCACTCCAGGAGCTCCT 1140
Qy 1141 GTAGTGACGAGAGCAGCTCAAGCATCAGCTGCGGGGACTCTCGGAAACCTCTGAG 1200
Db 1141 GTAGTGACGAGAGCAGCTCAAGCATCAGCTGCGGGGACTCTCGGAAACCTCTGAG 1200
Qy 1201 CTTGTTTCTTATTAATTAATTCAGCATAAACCCTTAAAAAATCTCTGAGGAACCTTTTCAAGAA 1260
Db 1201 CTTGTTTCTTATTAATTAATTCAGCATAAACCCTTAAAAAATCTCTGAGGAACCTTTTCAAGAA 1260
Qy 1261 ATTGATGGGGTGGGACACACAGCTACAGTGTGCTGGGACTAAGTCCCTACTCGGATAT 1320

|||||
1261 ATTGATGGGTGGCGACACCGCTCAGTGTGCTGGACCTAAGTCCCTACTCGATTAT 1320
Qy
1321 GAATTCAGGGTTGTTGCTGTCAATAAATTTGGGGGGGGCTCCAGCGAACTGTGCTA 1380
Db
1321 GAATTCAGGGTTGTTGCTGTCAATAAATTTGGGGGGGGCTCCAGCGAACTGTGCTA 1380
Qy
1381 ACACAAACCTCAGAGCAAGCAACATCCAGTGCCTCCGAGGAGATGTCCAGGCAAGATGTTG 1440
Db
1381 ACACAAACCTCAGAGCAAGCAACATCCAGTGCCTCCGAGGAGATGTCCAGGCAAGATGTTG 1440
Qy
1441 AGTTTCAGACCACTTTTGGTACAGTGAAGCACTCAAGAGCCAAATGACAGATCCAA 1500
Db
1441 AGTTTCAGACCACTTTTGGTACAGTGAAGCACTCAAGAGCCAAATGACAGATCCAA 1500
Qy
1501 GGATATAGAGTTTATATACAATGGATCCCACTCAACATGTCAACAACTGGATGAACAC 1560
Db
1501 GGATATAGAGTTTATATACAATGGATCCCACTCAACATGTCAACAACTGGATGAACAC 1560
Qy
1561 AATGTAGTGAAGCCAAATCACTACTATTGGGCAACTTAGTGCCCCAGAAACATATTCT 1620
Db
1561 AATGTAGTGAAGCCAAATCACTACTATTGGGCAACTTAGTGCCCCAGAAACATATTCT 1620
Qy
1621 GTCAAGTCTCGGCTTTTACCTCAATTTGGAGATGCTCCCTTTCAAGTGCATACAAGTC 1680
Db
1621 GTCAAGTCTCGGCTTTTACCTCAATTTGGAGATGCTCCCTTTCAAGTGCATACAAGTC 1680
Qy
1681 ATCACTCAGACAGAGTACAGGCGAGCCACTAACTTTCAAGCAGAACTGTGCTGAA 1740
Db
1681 ATCACTCAGACAGAGTACAGGCGAGCCACTAACTTTCAAGCAGAACTGTGCTGAA 1740
Qy
1741 ACAAGTATTTTGTCTCTTTGACACCTCCAGTTCAGATACCATTTGCCAATATGAAGTC 1800
Db
1741 ACAAGTATTTTGTCTCTTTGACACCTCCAGTTCAGATACCATTTGCCAATATGAAGTC 1800
Qy
1801 GTCTACAAGATGGGAGCATGGAGAGAGCAACGAATTTACATTTGAGCCAGGACATCA 1860
Db
1801 GTCTACAAGATGGGAGCATGGAGAGAGCAACGAATTTACATTTGAGCCAGGACATCA 1860
Qy
1861 TATAGGCTGCAAGGACTGAAACCAACAGCTTATATTTCCGCTGTGGCTGACGCTCC 1920
Db
1861 TATAGGCTGCAAGGACTGAAACCAACAGCTTATATTTCCGCTGTGGCTGACGCTCC 1920
Qy
1921 CCTAAGGCTGGGTGCTTCTACTGCAAAATATCAGCTAGAACCATGTCAAAGCCG 1980
Db
1921 CCTAAGGCTGGGTGCTTCTACTGCAAAATATCAGCTAGAACCATGTCAAAGCCG 1980
Qy
1981 TCAGTCTCTCCTCAAGACATTTAGTTGACAGCCCAAGTTCACCTAGTATTTTGTGAAGT 2040
Db
1981 TCAGTCTCTCCTCAAGACATTTAGTTGACAGCCCAAGTTCACCTAGTATTTTGTGAAGT 2040
Qy
2041 TGGCAACCTCCACAGTGGAAAAACAGAAATGGCAATTTACTGAAATCTCCATCAAGTAC 2100
Db
2041 TGGCAACCTCCACAGTGGAAAAACAGAAATGGCAATTTACTGAAATCTCCATCAAGTAC 2100
Qy
2101 ACTCAGTGGATGGGAAAGATGAACAGCTCAAGATTTTGGAAATTTCTTCGACACT 2160
Db
2101 ACTCAGTGGATGGGAAAGATGAACAGCTCAAGATTTTGGAAATTTCTTCGACACT 2160
Qy
2161 ACCAAATACCTTTTGGACAGCTGGAAAAATGGAATACCGGATCAGTGCACAGCC 2220
Db
2161 ACCAAATACCTTTTGGAAACAGCTGGAAAAATGGAATACCGGATCAGTGCACAGCC 2220
Qy
2221 CATACAGATGTCGGCCCTGGAGCTTTGTCGTTGATTTGCAATCGAAACCAATGAAGAT 2280
Db
2221 CATACAGATGTCGGCCCTGGAGCTTTGTCGTTGATTTGCAATCGAAACCAATGAAGAT 2280
Qy
2281 GTTCTAGTGGTCTCTCGAAAGTCCAGGTGAGGCTGTCAACTCAACATCTGTTAAA 2340
Db
2281 GTTCTAGTGGTCTCTCGAAAGTCCAGGTGAGGCTGTCAACTCAACATCTGTTAAA 2340
Qy
2341 GTCTCATGGCGCTCACCGCTGCCCAATAAAGCATGGCCAGATGAAGGATATCAGGTG 2400

Db
2341 GTCTCATGGCGCTCACCGCTGCCCAATAAAGCATGGCCAGATGAAGGATATCAGGTG 2400
Qy
2401 CATTATGTAGGATGGAATGTTGAGCCCAAGGGCCAGCCATGCTGAAAGATGTCATG 2460
Db
2401 CATTATGTAGGATGGAATGTTGAGCCCAAGGGCCAGCCATGCTGAAAGATGTCATG 2460
Qy
2461 CTGGCTGATGTCAGTGGGAATTTGATGATACCTACTGAAACATGACATGATCATTTCTGGG 2520
Db
2461 CTGGCTGATGTCAGTGGGAATTTGATGATACCTACTGAAACATGACATGATCATTTCTGGG 2520
Qy
2521 CTCACGCTGAAATCTTCTACTCCCTCAGCTCAGCTCAGCTCACAACCAAGAGATGTT 2580
Db
2521 CTCACGCTGAAATCTTCTACTCCCTCAGCTCAGCTCAGCTCACAACCAAGAGATGTT 2580
Qy
2581 GCTCGCAGCAAGCCCAAACTGTTGTCACACTGGGGCAGTTCCAGGGAACCTCCGCTT 2640
Db
2581 GCTCGCAGCAAGCCCAAACTGTTGTCACACTGGGGCAGTTCCAGGGAACCTCCGCTT 2640
Qy
2641 GTGATTTAAACCACTCAGATGAATACCTGCTCTTATTTAGTGGCACCCTCCGTTGGACACA 2700
Db
2641 GTGATTTAAACCACTCAGATGAATACCTGCTCTTATTTAGTGGCACCCTCCGTTGGACACA 2700
Qy
2701 TTTTGGACCTCTTTCAGGCTACCTCTAAATTTTGGCCGCAAGATATGAGAGCCACTTACT 2760
Db
2701 TTTTGGACCTCTTTCAGGCTACCTCTAAATTTTGGCCGCAAGATATGAGAGCCACTTACT 2760
Qy
2761 ACTCTTGGATTTCTGTAAGAAAGATCACTTTTACAGCTACACATCCACAGGGAGCA 2820
Db
2761 ACTCTTGGATTTCTGTAAGAAAGATCACTTTTACAGCTACACATCCACAGGGAGCA 2820
Qy
2821 TCATACGCTCTTCAGGCTCTCAGCCAGAAACAAAGTGGGCTTTGGGGAGGAGATGGTGAAG 2880
Db
2821 TCATACGCTCTTCAGGCTCTCAGCCAGAAACAAAGTGGGCTTTGGGGAGGAGATGGTGAAG 2880
Qy
2881 GAGATTTCCATTTCCAGAGAAAGTACCAACTGGATTCCTTCAAAACCTTCACTCAGAAAGC 2940
Db
2881 GAGATTTCCATTTCCAGAGAAAGTACCAACTGGATTCCTTCAAAACCTTCACTCAGAAAGC 2940
Qy
2941 ACCACTTCAACCTCCGTCAGTTTATCTTGGCAACCACTGTCTCGCAGCAGAGAAATGGC 3000
Db
2941 ACCACTTCAACCTCCGTCAGTTTATCTTGGCAACCACTGTCTCGCAGCAGAGAAATGGC 3000
Qy
3001 ATTATCAACCAAGTATACCTTCTTTATAGGGAATCAACATCCCTTCTCCCGATGGAG 3060
Db
3001 ATTATCAACCAAGTATACCTTCTTTATAGGGAATCAACATCCCTTCTCCCGATGGAG 3060
Qy
3061 CAGCTTATTTCTCCAGCTGACCACTATGACACTCACTGGCTTAAACCCAGATACACA 3120
Db
3061 CAGCTTATTTCTCCAGCTGACCACTATGACACTCACTGGCTTAAACCCAGATACACA 3120
Qy
3121 TAGATGTAAGTACGTGCTCATACGAGCAAGGGCCCGGCCATATAGTCCCGATGTC 3180
Db
3121 TAGATGTAAGTACGTGCTCATACGAGCAAAAGGGCCCGGCCATATAGTCCCGATGTC 3180
Qy
3181 CAGTTCAGGACACTGCTGCTGGATCAAGTGTGTTGCAAAAAATTTTCAATGTCAAAAGCAGTA 3240
Db
3181 CAGTTTCAAGGACACTGCTGCTGGATCAAGTGTGTTGCAAAAAATTTTCAATGTCAAAAGCAGTA 3240
Qy
3241 ATGAAGACTTCCGTTGCTGTCTTGGGAGATTCAGAGAAATTTATACTCCGCTATGCT 3300
Db
3241 ATGAAGACTTCCGTTGCTGTCTTGGGAGATTCAGAGAAATTTATACTCCGCTATGCT 3300
Qy
3301 TTCAAAATTTCTTATGATGAGGAAAAATGTTAGAAAGTGGATGGCCGAGCCACACAG 3360
Db
3301 TTCAAAATTTCTTATGATGAGGAAAAATGTTAGAAAGTGGATGGCCGAGCCACACAG 3360
Qy
3361 AAGTTAATTTGTCAACCTGAAAGCTGAGAAATCATATTTTGTGCTGACAAATCGTGA 3420
Db
3361 AAGTTAATTTGTCAACCTGAAAGCTGAGAAATCATATTTTGTGCTGACAAATCGTGA 3420
Qy
3421 AACAGTCTGTGGGCTGCAGCAGGCTCAAGGCAAGAGCTGACAGATGTTATACGT 3480
Db
3421 AACAGTCTGTGGGCTGCAGCAGGCTCAAGGCAAGAGCTGACAGATGTTATACGT 3480

3481 QY ACCAAGCCTGCTTCAATTTGGGAAGACCACAACTTTGGATGGCATGATTTACTGTGCAACTGCT 3540
3481 Db ACCAAGCCTGCTTCAATTTGGGAAGACCACAACTTTGGATGGCATGATTTACTGTGCAACTGCT 3540
3541 QY GAAGTACCTGCAAAATGAGAATATAAAGGTTACTACATAAATAATTTGTGCTTTGAAGAA 3600
3541 Db GAAGTACCTGCAAAATGAGAATATAAAGGTTACTACATAAATAATTTGTGCTTTGAAGAA 3600
3601 QY TCTCGCGGAAATTTTATCAAGCCATGGGAGAGTCCAGATGAAATGGAATTAGATGAGCTG 3660
3601 Db TCTCGCGGAAATTTTATCAAGCCATGGGAGAGTCCAGATGAAATGGAATTAGATGAGCTG 3660
3661 QY CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGAGAGAGTTGAAATTAAG 3720
3661 Db CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGAGAGAGTTGAAATTAAG 3720
3721 QY CCATATATTTGCGGCTCACCTTTGATGCTCTTCCACCTGAGTTTCCCTCGGGGATGACAAG 3780
3721 Db CCATATATTTGCGGCTCACCTTTGATGCTCTTCCACCTGAGTTTCCCTCGGGGATGACAAG 3780
3781 QY CATTATGTTGATTTTACAAAACAGCAACTCCAAAGTGGTCAAGATATGTTCTTTCTG 3840
3781 Db CATTATGTTGATTTTACAAAACAGCAACTCCAAAGTGGTCAAGATATGTTCTTTCTG 3840
3841 QY TTAGCAGTAATGGAACATGCGAGGCTTAAGATGTATGCAACGAGCCCTTATCTCGACCCC 3900
3841 Db TTAGCAGTAATGGAACATGCGAGGCTTAAGATGTATGCAACGAGCCCTTATCTCGACCCC 3900
3901 QY GTGGTGTCAATGGAATCTGGATCCGAGCCCAATCACGGATGAAGAGAGGCTTGATCTGG 3960
3901 Db GTGGTGTCAATGGAATCTGGATCCGAGCCCAATCACGGATGAAGAGAGGCTTGATCTGG 3960
3961 QY GTTGTAGTCTGCTGCTTCCAGTGTCTTTATCATCTGCAATTTGTCATTTGCTTCTT 4020
3961 Db GTTGTAGTCTGCTGCTTCCAGTGTCTTTATCATCTGCAATTTGTCATTTGCTTCTT 4020
4021 QY TATATAAGGAGAGGCGCAGGCTCCGACTCTAGAAAAAGCAGCATACCCGAAACAATAAGGAG 4080
4021 Db TATATAAGGAGAGGCGCAGGCTCCGACTCTAGAAAAAGCAGCATACCCGAAACAATAAGGAG 4080
4081 QY ATCCCTTCCACACCCACAGCCCTGTAGACTGAGGCGCTTAACTTTCAACACCG 4140
4081 Db ATCCCTTCCACACCCACAGCCCTGTAGACTGAGGCGCTTAACTTTCAACACCG 4140
4141 QY GGTATGGCTAGCCATCTCCCAATACCCATCTTTGAACTTTGCAAGATG 4200
4141 Db GGTATGGCTAGCCATCTCCCAATACCCATCTTTGAACTTTGCAAGATG 4200
4201 QY AAAGCAAAATGCACTTGAAGTTTCCAGGAATATGATCAATTTGACCCCTGGCCAGCAG 4260
4201 Db AAAGCAAAATGCACTTGAAGTTTCCAGGAATATGATCAATTTGACCCCTGGCCAGCAG 4260
4261 QY TTCACCTTGGGAACATTTCAAATCTTGAAGTAAACCAAGATAGATAGATAGATAGATAGAT 4320
4261 Db TTCACCTTGGGAACATTTCAAATCTTGAAGTAAACCAAGATAGATAGATAGATAGATAGAT 4320
4321 QY ATCGCATATGATCATTTCCCGGTTCTCTATCAGCTATAGAGGATCCAGGAGTGAC 4380
4321 Db ATCGCATATGATCATTTCCCGGTTCTCTATCAGCTATAGAGGATCCAGGAGTGAC 4380
4381 QY TATGTGAATGCAACTACATAGATGGGTATAGGAAGCAAAATGCTTATTTGCAACACAG 4440
4381 Db TATGTGAATGCAACTACATAGATGGGTATAGGAAGCAAAATGCTTATTTGCAACACAG 4440
4441 QY GGATCTCTCCCGAAACATTTGGGACTTTTGGAGATGATATGGGAACAACGAGTGCC 4500
4441 Db GGATCTCTCCCGAAACATTTGGGACTTTTGGAGATGATATGGGAACAACGAGTGCC 4500
4501 QY ACAGTTGTATGATGACAAAACCTAGAAAGAGATCAAGGGTGAAGTGTGACCATTTGG 4560
4501 Db ACAGTTGTATGATGACAAAACCTAGAAAGAGATCAAGGGTGAAGTGTGACCATTTGG 4560

4561 QY CCTAGCAGAGCAGAGAAACCCACGGAAGTCTGTTCAAGTAAACGCTGCTTGAATCTGTGAG 4620
4561 Db CCTAGCAGAGCAGAGAAACCCACGGAAGTCTGTTCAAGTAAACGCTGCTTGAATCTGTGAG 4620
4621 QY CTGGCCACATATTTGTTTCAAGCAATTTGTCATCTTTTACAGAAATGGTTCAAGTGAGAAGA 4680
4621 Db CTGGCCACATATTTGTTTCAAGCAATTTGTCATCTTTTACAGAAATGGTTCAAGTGAGAAGA 4680
4681 QY GAAGTGAGACAAATTTCCAGTTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4740
4681 Db GAAGTGAGACAAATTTCCAGTTTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4740
4741 QY CCTTTTCTAGCTTTTCTTACGTAGAGTCAAAACCTGTAACCTTCCCGATGCTGCTGCTGCTGCT 4800
4741 Db CCTTTTCTAGCTTTTCTTACGTAGAGTCAAAACCTGTAACCTTCCCGATGCTGCTGCTGCTGCT 4800
4801 QY GTTGTGCACTGCACTGCGGAGTTGGCGGAGTGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGT 4860
4801 Db GTTGTGCACTGCACTGCGGAGTTGGCGGAGTGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGT 4860
4861 QY TTAGAAAGAAATAAGCATGAAATAAATCTGTAGATATTTATGGCCATGTAACTTTTAAATGAGA 4920
4861 Db TTAGAAAGAAATAAGCATGAAATAAATCTGTAGATATTTATGGCCATGTAACTTTTAAATGAGA 4920
4921 QY GCCCAGAGGAATATATGTTTCAAAACAGAGCAATATATCTTTATCCATGATGCACTG 4980
4921 Db GCCCAGAGGAATATATGTTTCAAAACAGAGCAATATATCTTTATCCATGATGCACTG 4980
4981 QY TTAGAAGCAGTACTTGTGGAATACCGAGTGGCAGTACGAACTTGTATGCTTACCTTACAT 5040
4981 Db TTAGAAGCAGTACTTGTGGAATACCGAGTGGCAGTACGAACTTGTATGCTTACCTTACAT 5040
5041 QY CAGAAGCTGACACAAATAGAAACGGAGAGATGTCAAGGAATGGAGCTCGAAATTTAAG 5100
5041 Db CAGAAGCTGACACAAATAGAAACGGAGAGATGTCAAGGAATGGAGCTCGAAATTTAAG 5100
5101 QY CGTCTAGCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCATCTTCCATGTAAT 5160
5101 Db CGTCTAGCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCATCTTCCATGTAAT 5160
5161 QY AAATTTCAAAATCGCTTGTAAATATTTATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 5220
5161 Db AAATTTCAAAATCGCTTGTAAATATTTATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 5220
5221 QY CCTATCGGTGAGTAGAAGATCTGATTTACATCAATGCAATGCAATGCAATGCAATGCAATGCAAT 5280
5221 Db CCTATCGGTGAGTAGAAGATCTGATTTACATCAATGCAATGCAATGCAATGCAATGCAATGCAAT 5280
5281 QY CAACAGAAAGCTACATGCTTACCCAGGCGCTTGGCAGAGACCACTGGAAGACTTCTGG 5340
5281 Db CAACAGAAAGCTACATGCTTACCCAGGCGCTTGGCAGAGACCACTGGAAGACTTCTGG 5340
5341 QY CGGATGCTCTGGGAACAAATTTCCACCATTAGTTGTGATGCTCACCAGCTCGGTTGAAATG 5400
5341 Db CGGATGCTCTGGGAACAAATTTCCACCATTAGTTGTGATGCTCACCAGCTCGGTTGAAATG 5400
5401 QY GGCAAGAGAAATGTCACCAATATCTGGCCAGCAAGCGGTCTGCAAGATACCAAGTACTTT 5460
5401 Db GGCAAGAGAAATGTCACCAATATCTGGCCAGCAAGCGGTCTGCAAGATACCAAGTACTTT 5460
5461 QY GTTGTAGATCCATGCTGCTGAGTACCAATGCTGCAAGTATATCTTAAGGAAATTAAGGTC 5520
5461 Db GTTGTAGATCCATGCTGCTGAGTACCAATGCTGCAAGTATATCTTAAGGAAATTAAGGTC 5520
5521 QY ACAGATCCAGGAGCGCCAGTCCCGAACAGTAGTAGGAGTTCACAGTCTGCTGCTGCTGCTGCTGCT 5580
5521 Db ACAGATCCAGGAGCGCCAGTCCCGAACAGTAGTAGGAGTTCACAGTCTGCTGCTGCTGCTGCTGCT 5580
5581 QY GAGCAAGAGTGCACAAAGTCCGGAAGAGATTTTATGACTTTCATCGGCAAGTCCATATAA 5640
5581 Db GAGCAAGAGTGCACAAAGTCCGGAAGAGATTTTATGACTTTCATCGGCAAGTCCATATAA 5640
5641 QY ACAAAGAAACAGTTTGTGCGCAAGATGGACCCATTTTCACTGTCATTTGACGGCGGCTTGG 5700

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 19:22:54 ; Search time 217 Seconds
(without alignments)
3407.769 Million cell updates/sec

Title: US-10-772-636-64

Perfect score: 10042

Sequence: 1 MVHVARLLLLLLFFLRLTDA.....YQFSYRAALEYLGLSPDHYAT 1912

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq 16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10042	100.0	1912	8	ADR40183 Human pro
2	7369.5	73.4	1949	7	ADE57117 Human pro
3	7369.5	73.4	1949	7	ADE57121 Human pro
4	7369.5	73.4	1949	7	ADD47015 Human pro
5	7369.5	73.4	1949	7	ADD47015 Human pro
6	7359.5	73.3	1948	7	ADD18742 Human dis
7	7245	72.1	1911	2	AAR71726 Human PTP
8	7245	72.1	1911	2	AAR72225 Human pro
9	7245	72.1	1911	2	AAW94027 Human pro
10	7245	72.1	1911	4	AAU01459 Human pro
11	7110	70.8	1904	5	ABBS7100 Mouse isc
12	7037	70.1	1907	4	AAU14143 Human nov
13	7032	70.0	1897	3	AAAY81785 Human pro
14	7032	70.0	1897	3	AAAY81785 Human pro
15	7032	70.0	1897	3	AAAB19712 Human pro
16	7032	70.0	1897	7	ADD18740 Human dis
17	7032	70.0	1897	8	ADJ33670 Human leu
18	7032	70.0	1907	8	ADP18674 Human pro
19	6977.5	69.5	1863	7	ADDA46989 Rat Prote
20	6823.5	66.0	1786	6	AAE37971 Human kin
21	6308.5	62.8	1254	8	ADN02662 Liver dis
22	6251.5	62.3	1291	2	AAAR75201 Tyrosine
23	7245.5	57.2	1501	2	AAAR72858 Rat recep
24	5559.5	55.4	1495	5	ABBS7380 Rat mucoc
25	5559.5	55.4	1496	7	ADE57115 Rat Prote

26	5559.5	55.4	1496	7	ADE57119 Rat Prote
27	5559.5	55.4	1496	7	ADD47013 Rat Prote
28	5559.5	55.4	1496	7	ADD47017 Rat Prote
29	5333	53.1	1477	8	ADJ33671 Mouse leu
30	4865	48.4	2037	4	ABB71928 Drosophil
31	4664.5	46.4	1366	8	ADO66041 Novel hum
32	3655	36.4	960	8	ADR08858 Human pro
33	3027.5	30.1	647	4	AAW23746 Human EST
34	3027.5	30.1	647	4	AAU14379 Human nov
35	3027.5	30.1	647	8	ADH80697 Human pro
36	2965.5	29.5	1194	6	ABP97679 Amino aci
37	2926	29.1	607	3	AAAY81783 Human pro
38	2926	29.1	607	3	AAAY56098 LAR tyros
39	2238	22.3	1585	8	ADN22916 Bacterial
40	2238	22.3	1585	8	ADN22917 Bacterial
41	2166	21.6	442	3	AAAB56372 Human pro
42	2024	20.2	608	8	ADR09270 Human pro
43	1957.5	19.5	400	2	AAAR75203 Tyrosine
44	1814	18.1	344	7	ADJ69864 Human hea
45	1587	15.8	793	2	AAAR20743 Murine re

ALIGNMENTS

RESULT 1

ADR40183
ID ADR40183 standard; protein; 1912 AA.

XX ADR40183;

AC AC (first entry)

DT 18-NOV-2004

XX Human protein tyrosine phosphatase delta (R-PTP-delta) (965) protein.

DE haematological; cytostatic; erythroid; anaemia; erythrocytosis;

XX bone marrow; leukaemia; platelet; thrombocytopenia; thrombosis; B-cell;

KW T-cells; neutropenia; gene therapy; human;

KW protein tyrosine phosphatase delta; R-PTP-delta; enzyme.

XX Homo sapiens.

OS WO2004072242-A2.

XX 26-AUG-2004.

PD 05-FEB-2004; 2004WO-US003417.

XX 05-FEB-2003; 2003US-0445241P.

XX 18-FEB-2003; 2003US-044389P.

XX 20-MAR-2003; 2003US-0456320P.

XX 03-APR-2003; 2003US-0460279P.

XX 28-APR-2003; 2003US-0465924P.

XX 13-MAY-2003; 2003US-0470052P.

XX 26-AUG-2003; 2003US-0498106P.

XX 04-SEP-2003; 2003US-0500179P.

XX 15-SEP-2003; 2003US-0502909P.

XX 10-OCT-2003; 2003US-0510351P.

XX 17-OCT-2003; 2003US-0512380P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Kelly LM, Carroll JM, Farlow D, Healy A;

XX WPI; 2004-625850/60.

XX N-PSDB; ADR40182.

XX Identifying a compound capable of treating a hematological disorder

PT comprises combining a compound to be tested with a polypeptide related

PT with the disorder under conditions suitable for binding of the test

XX compound to the polypeptide.

XX Claim 1; SEQ ID NO 64; 321pp; English.

Db 1861 LERMRYEGVVDIFQTVKMLRTQRPAMVQTEDQYQFSYRAALEYLGSFPHYAT 1912

RESULT 2

AD557117

ID ADE57117 standard; protein; 1949 AA.

XX AC ADE57117;

XX DT 29-JAN-2004 (first entry)

XX DE Human Protein U35234, SEQ ID NO 2977.

XX KW Human; pain; neuronal tissue; gene therapy;

XX KW spinal segmental nerve injury; chronic constriction injury; CCI;

XX KW spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GHEO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX DR GENBANK; U35234.

XX PT New composition comprising two or more isolated polypeptides, useful for

XX PT preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat

XX CC or human polynucleotides or a polynucleotide which represents a fragment,

XX CC derivative or allelic variation of the nucleic acid sequence. Also

XX CC claimed are a vector comprising the novel polynucleotide, a host cell

XX CC comprising the vector, a method for identifying a nucleotide sequence

XX CC which is differentially regulated in an animal subjected to pain and a

XX CC kit to perform the method, an array, a method for identifying an agent

XX CC that increases or decreases the expression of the polynucleotide sequence

XX CC that is differentially expressed in neuronal tissue of a first animal

XX CC subjected to pain, a method for identifying a compound which regulates

XX CC the expression of a polynucleotide sequence which is differentially

XX CC expressed in an animal subjected to pain, a method for identifying a

XX CC compound that regulates the activity of one or more of the

XX CC polynucleotides, a method for producing a pharmaceutical composition, a

XX CC method for identifying a compound or small molecule that regulates the

XX CC activity in an animal of one or more of the polypeptides given in the

XX CC specification, a method for identifying a compound useful in treating

XX CC pain and a pharmaceutical composition comprising the one or more

XX CC polypeptides or their antibodies. The polynucleotide or the compound that

XX CC modulates its activity is useful for preparing a medicament for treating

XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

XX CC therapy). The sequence presented is a human protein (shown in Table 2 of

XX CC the specification) which is differentially expressed during pain. Note:

XX CC The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic form directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1949 AA;

Query Match 73.4%; Score 7369.5; DB 7; Length 1949;

Best Local Similarity 71.9%; Pred. No. 0;

Matches 1384; Conservative 224; Mismatches 281; Indels 35; Gaps 11;

Qy 20 AETPPRTPRTPVDTQVSGGVASFICQATGDPRPKIWMNKKGVSNQRFEVIEFDGSG 79

Db 29 AEEPFRPIKEPKDQIGVSGGVASFVCOATGDPKPRVTWNKKGVSNQRFEVIEFDGSG 88

Qy 80 SVLRIOPLRTPRDEAIYECVANNVGEISVSTLTLVLREDOI PRGPTIDMGQPKVVER 139

Db 89 AVLRIQPLRTPRDEAIYECVANNVGEISVSTLTLVLREDOI PRGPTIDMGQPKVVER 148

Qy 140 TRTATMLCAASGNPDPEITWFKDFLPVDTSNNGRIKQLRSEISIGGTPIRGALQIEOSE 199

Db 149 TRTATMLCAASGNPDPEITWFKDFLPVDPASNGRIKQLRSEISIGGTPIRGALQIEOSE 208

Qy 200 SDQGYECVATNSAGTRYSAANLYVRELREVRVRPRFSIPTNNHEIMPGGSVNITCVA 259

Db 209 TDQGYECVATNSAGTRYSAANLYVRELREVRVRPRFSIPTNNHEIMPGGSVNITCVA 268

Qy 260 VGSFMPYVKKMLCAEDLTPEDDDMPGRNVLELNDVRSQSANLYTCVAMSTGLVIAIAQITV 319

Db 269 VGSFMPYVKKMLCAEDLTPEDDDMPGRNVLELNDVRSQSANLYTCVAMSTGLVIAIAQITV 328

Qy 320 KALPKDPGTPVVTTESTATSTLTWDSGNPEPVSYIIQHKPKNSEELYKEIDGVATTRY 379

Db 329 KSLPKAPGTPVVTTESTATSTLTWDSGNPEPVSYIIQHKPKNSEELYKEIDGVATTRY 388

Qy 380 VAGLSPYSYEFYRVAANNIGRPPSEPVLTQTSQAAPSAPRDVQARMLSSITLTVQWK 439

Db 389 IGLSPNSEYEIWSAVNSIGQGPSPSESVVTRTGEQAPASAPRVQARMLSATMTIQQWE 448

Qy 440 EPEEPNGOIQGYRYVTMDPTQHVNNWKNVADSOITITIGNLVPOKTYSVKVLAFISIG 499

Db 449 EPVEPNGLIRGYRYVTMEPEHPVGNWQKHNDVDSLLTTVGSLEDETYTVRLAFISVG 508

Qy 500 DGPLSSDIQVITOTGVPGOPLPNFKAPSESTSLTSLWTPRSDTIANYELVYKDGHEE 559

Db 509 DGPLSDPIQVKTQGVPGOPMNRARESETSTLTSWSPRQESIYKELFFRGDGHRE 568

Qy 560 QRITIEPGTSYRLQGLKPNLSLYYFRLAARSPQGLGASTABISARTMQSPKSPAPQDISCT 619

Db 569 VGRTPDPTTSYVVEDLKPNTYAFRLAARSPQGLGASTABISARTMQSPKSPAPQDKVCV 628

Qy 620 SPSTSLVSWQPPVPEKONGIITEYSIKYTAVDGDDKPHILGILGPSOTTKYLLEOLEK 679

Db 629 SVRSTAILVSWRPPPEPTHNGALVGYSVRYRPLGSEDPKPEVNGIPPTTQTLLEALEK 688

Qy 680 WTEYRITVTAHTDVGPPESLSVLIRTNEDVPSGPPKVEAVEAVNSTSVKSVSRPVNPK 739

Db 689 WTQYRITVTAHTDVGPPESLSVLIRTNEDVPSGPPKVEAVEAVNSTSVKSVSRPVNPK 748

Qy 740 QHQIIRGYQVHYVRMENGEPKQOPMLKDVMLADAQWEDDTTEHDMIISGLQPEYSYSLT 799

Db 749 QHQIIRGYQVHYVRMENGEPKQOPMLKDVMLADAQWEDDTTEHDMIISGLQPEYSYSLT 808

Qy 800 VTATYTKGAGSKPKLVSTTGAVPKPRLVNHNTQMTALIQWHPVDTFF-GPLOGYRL 858

Db 809 VAAVTYTKGAGSKPKLVSTTGAVPKPRLVNHNTQMTALIQWHPVDTFF-GPLOGYRL 868

Qy 859 KGRKDMELTTLTLESEKEDHFTATDIHKGASVVFELSAKNKVGCEGMVKETSIPEEVP 918

Db 869 QFGREDSTPLATLEFPSPSDRYTAGSVHKGATYVFLAARSRGGLGEEAAVLSIPEDT 928

Qy 919 TQFPQNLHSEGTST-SVLSWQPPVLAERNGLIITKYLTYRDLINPLLPMEOLIVP--- 974

Db 929 RGHPOLEAAGNASAGTVLLRMLPPVPAERNGAIKVTYVAVREAG-ALGPARETELPA 987

Qy 975 ----ADTTMTLTLGLKPTDVTYKVRATSKGPGYSPSVQFRTLFPVDQVFAKNPHVAVMK 1031

Db 988 EPGAENAVTLQGLKPDYDLQVRAHTRRGPSPSPVRYRTFLRDQVSPKPKVIMK 1047

Qy 1032 TSVLLSWEIPEVNSAMPEKILYDDGKVEEDVGRATOKLIVNLKPEKSVFVLTNRGNS 1091

Db 1048 TSVLLSWEFPDNNSTPTFYKIQNG--LTLVDGRGRTTKLILHLKPHFTFNFVLTNRGSS 1105
 Qy 1092 AGGLQHVRTAKTADVLRTKPAFIGNKGLDMITVQLPEYPAENIKGYVIIIVPLKKSQ 1151
 Db 1106 LGGQQVTVTAWTAFNLNGKSPVAPKPDAGFIMVLPDQGSVPVQSYFIVWVPLKSR 1165
 Qy 1152 -GKFIKWPSPDEMEDELLEKISR-KRRSIRYGREVEL-KPYIAAHFDVLPTFTLGDD 1208
 Db 1166 GGQFLTLPLGSPEDWLEELIQDISRLQRTVRHSRQLEVPYIAARFSLVLPTEHPGQ 1225
 Qy 1209 KHGGFTNKQLSQSQEYVFFVLAWMEHAESKMVATSPSPVSMOLDPOPIITDEBGLI 1268
 Db 1226 KQYGGFNRGLPEGRHYLVFLAVLQKSE-PTFAASPFSDPFLQNDPDPQFIVDGEGLI 1284
 Qy 1269 WVGSPVLAUVFIIICIVAILLYKR---KRAESDSKSSIPNNKEIPSHHPTDPVLELRL 1324
 Db 1285 WVGSPVLAUVFIIICIVAILLYKKNKPSKRDSEPRTKCLNNADLAPHPKDPVEMRI 1344
 Qy 1325 NFQTP-----GMASHPPPIPILELADHIERLKANDNLKFSQYESIDPGQ 1368
 Db 1345 NFQTPDGLRSLRPEPGFHFESMLSHPPPIADMAEHTERLKANDSLKLSQYESIDPGQ 1404
 Qy 1369 QFTWEHNLVKNPKRYANVIAVDSRVLLSAIEGIPGSDYVNAVYIDGYRQKNAYIAT 1428
 Db 1405 QFTWEHNLVKNPKRYANVIAVDFRVLQPIETIGMSDYINANYVDGYRRQNAVYIAT 1464
 Qy 1429 QGSLPETFGDFWRMIFQRSATVVMTKLEERSRVKCDQYWPSSRGTEHGLVQVTLDDTV 1488
 Db 1465 QGSLPETFGDFWRMIFQRSATVVMTKLEERSRVKCDQYWPSSRGTEHGLVQVTLDDTV 1524
 Qy 1489 ELATYCVRTFALYKNGSSKREVRQFTAWPDHGVPEHPTFPLAFLRRVKTNCNPDAGP 1548
 Db 1525 ELATYCVRTFSLHNGSSKREVRQFTAWPDHGVPEHPTFPLAFLRRVKTNCNPDAGP 1584
 Qy 1549 MVVHCSAGVGTGCFIIVDAMLERIKHEKTVYIGHVTLMAQRNVYVQTEQYIFIHDA 1608
 Db 1585 IIVHCSAGVGTGCFIIVDAMLERIKHEKTVYIGHVTLMAQRNVYVQTEQYIFIHDA 1644
 Qy 1609 LLEAVTCNTEVPARNLYAYIKLQIETGENTVGMLEFKRLASSKAHSTRISANLPC 1668
 Db 1645 LLEAVTCNTEVPARNLYAYIKLQIETGENTVGMLEFKRLASSKAHSTRISANLPC 1704
 Qy 1669 NKFNRLVNIIMPYSTVCLQIRGVGSDYINASFIDGYRQKQKAYIATQGLAETTFD 1728
 Db 1705 NKFNRLVNIIMPYSTVCLQIRGVGSDYINASFIDGYRQKQKAYIATQGLAETTFD 1764
 Qy 1729 WRMLWEHNSITVWMLTKLREMGREKCHQYWPASRSARYQYFVVDPMAYNMPQYILREFK 1788
 Db 1765 WRMLWEHNSITVWMLTKLREMGREKCHQYWPASRSARYQYFVVDPMAYNMPQYILREFK 1824
 Qy 1789 VTDARDGQSRVROFQFTDPEQGVKSGGFDIFIGVHKTKEQFGQDGPISVHCSAGV 1848
 Db 1825 VTDARDGQSRVROFQFTDPEQGVKSGGFDIFIGVHKTKEQFGQDGPISVHCSAGV 1884
 Qy 1849 GRTGVFTLSIVLERMYEGVVDIFQTVKMLRTQRPAMVQTEQYQFSYRAALEYLGSPD 1908
 Db 1885 GRTGVFTLSIVLERMYEGVVDIFQTVKMLRTQRPAMVQTEQYQFSYRAALEYLGSPD 1944
 Qy 1909 HYAT 1912
 Db 1945 HYAT 1948

RESULT 3
 ADE57121
 ID ADE57121 standard; protein; 1949 AA.
 AC ADE57121;
 XX
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Protein U35234, SEQ ID NO 2981.
 XX

KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 XX spared nerve injury; SNI; Chung.
 OS Homo sapiens.
 PN WO2003016475-A2.
 XX 27-FEB-2003.
 PD
 XX 14-AUG-2002; 2002WO-US025765.
 PF
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 PR
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PA
 XX
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI: 2003-268312/26.
 DR GENBANK; U35234.
 DR
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 XX preparing a medicament for treating pain in an animal.
 PT
 XX
 PS Claim 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 1949 AA;
 SQ

Query Match 73.4%; Score 7369.5; DB 7; Length 1949;
 Best Local Similarity 71.9%; Pred. No. 0;
 Matches 1384; Conservative 224; Mismatches 281; Indels 35; Gaps 11;

Qy 20 AETPRFTPTVDQTVSGVASFICQATGDRPKIVNKKGVSNORFIEVDFDGS 79
 Db 29 ABEPRFTKPKDQIGVSGVASFVQATGDRPKIVNKKGVSNORFIEVDFDGS 88
 Qy 80 SVLRIOPLRTPRDEAIECVASNNVGEISVSTRITVLRDQIPRGFTIDMGPKLVKVR 139
 Db 89 AVLRIOPLRTPRDEAIECVASNNVGEISVSTRITVLRDQIPRGFTIDMGPKLVKVR 148
 Qy 140 TTTATWMLCAASGNPDPEITWFKOPITVDTSTNNNGRIKQLRSSEIGTPIRGALQIESSE 199

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 19:37:10 ; Search time 52 Seconds
(without alignments)
2744.789 Million cell updates/sec

Title: US-10-772-636-64
Perfect score: 10042
Sequence: 1 MWHVRLLLLLLTFFLRDTA.....YQFSYRAALEYLGSFPHYAT 1912

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS COMB pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7245	72.1	1911	1	US-08-348-006B-5
2	7245	72.1	1911	2	Sequence 5, Appli
3	7245	72.1	1911	3	Sequence 5, Appli
4	7245	72.1	1911	5	Sequence 5, Appli
5	5745.5	57.2	1501	2	US-08-447-464-3
6	5745.5	57.2	1501	2	US-08-716-679-3
7	2646	26.3	538	4	US-09-743-492A-9
8	1587	15.8	793	1	US-08-015-985-3
9	1587	15.8	793	1	US-09-280-597-3
10	1565.5	15.6	802	1	US-08-015-985-1
11	1565.5	15.6	802	4	US-09-280-597-1
12	1565.5	15.6	807	4	US-09-949-016-7356
13	1542	15.4	1452	2	US-08-449-644-8
14	1541	15.3	1452	2	US-08-087-244A-8
15	1541	15.3	1452	2	US-08-652-971-4
16	1541	15.3	1452	2	US-08-991-258A-4
17	1541	15.3	1452	2	US-08-769-399-4
18	1541	15.3	1452	3	US-08-991-953A-4
19	1502.5	15.0	1439	2	US-08-449-644-2
20	1502.5	15.0	1439	2	US-08-087-244A-2
21	1482.5	14.8	1457	2	US-08-652-971-3
22	1482.5	14.8	1457	2	US-08-449-644-1
23	1482.5	14.8	1457	2	US-08-087-244A-1
24	1482.5	14.8	1457	2	US-08-991-258A-3
25	1482.5	14.8	1457	2	US-08-769-399-3
26	1482.5	14.8	1457	3	US-08-991-953A-3
27	1455	14.5	699	1	US-08-348-006B-7

28 1455 14.5 699 2 US-08-800-825A-7 Sequence 7, Appli
29 1455 14.5 699 3 US-09-158-657-7 Sequence 7, Appli
30 1450.5 14.4 1075 4 US-09-949-016-8308 Sequence 8308, Ap
31 1415 14.1 296 1 US-10-374-539-3 Sequence 3, Appli
32 1384 13.8 289 1 US-08-036-210-13 Sequence 13, Appli
33 1384 13.8 289 2 US-08-449-609-13 Sequence 13, Appli
34 1384 13.8 289 4 US-09-361-096A-13 Sequence 13, Appli
35 1381.5 13.8 560 4 US-09-949-016-10786 Sequence 10786, A
36 1325 13.2 1436 2 US-08-652-971-2 Sequence 2, Appli
37 1325 13.2 1436 2 US-08-991-258A-2 Sequence 2, Appli
38 1325 13.2 1436 2 US-08-769-399-2 Sequence 2, Appli
39 1325 13.2 1436 3 US-08-991-953A-2 Sequence 2, Appli
40 1291.5 12.9 1445 1 US-08-015-986A-2 Sequence 2, Appli
41 1291.5 12.9 1445 2 US-08-446-363-2 Sequence 2, Appli
42 1286 12.8 1442 1 US-08-015-986A-3 Sequence 3, Appli
43 1286 12.8 1442 2 US-08-446-363-3 Sequence 3, Appli
44 1250 12.4 245 2 US-08-685-992-26 Sequence 26, Appli
45 1250 12.4 245 2 US-09-144-925-26 Sequence 26, Appli

ALIGNMENTS

RESULT 1
US-08-348-006B-5
; Sequence 5, Application US/08348006B
; Patent No. 5658756
; GENERAL INFORMATION:
; APPLICANT: RODAN, GIDEON A.
; APPLICANT: SCHMIDT, AZRIEL
; APPLICANT: RUTLEDGE, SU JANE
; TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: J. MARK HAND
; STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,006B
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/122,032
; FILING DATE: 14-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: HAND, J., MARK
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 189921A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-3905
; TELEFAX: 908-594-4720
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1911 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-348-006B-5

Query Match 72.1%; Score 7245; DB 1; Length 1911;
Best Local Similarity 71.7%; Pred. No. 0;
Matches 1369; Conservative 219; Mismatches 279; Indels 42; Gaps 14;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 19:41:01 ; Search time 220 Seconds
(without alignments)
3352.912 Million cell updates/sec

Title: US-10-772-636-64
Perfect score: 10042
Sequence: 1 MVHVARLLLLLLLLFFLRTDA.....YQFSYRAALYLGFSDFHYAT 1912

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Seatched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCN_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCNUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10042	100.0	1912	17	US-10-772-636-64
2	7369.5	73.4	1948	9	US-09-808-602-55
3	7369.5	73.4	1948	10	US-09-800-198-45
4	7037	70.1	1907	15	US-10-291-265-250
5	7032	70.0	1897	17	US-10-482-023-52
6	7032	70.0	1897	18	US-10-712-892A-34
7	5794.5	57.7	1502	9	US-09-808-602-54
8	5794.5	57.7	1502	10	US-09-800-198-44
9	5559.5	55.4	1495	15	US-10-258-666-12
10	4877	48.6	2029	15	US-10-087-684-38
11	4877	48.6	2029	15	US-10-218-779-39

12	4865	48.4	2037	15	US-10-087-684-39	Sequence 39, Appl
13	4865	48.4	2037	15	US-10-218-779-39	Sequence 39, Appl
14	3027.5	30.1	647	15	US-10-291-265-722	Sequence 722, App
15	2965.5	29.5	1194	14	US-10-191-029-10	Sequence 10, Appl
16	2238	22.3	1585	15	US-10-369-493-5569	Sequence 5569, Ap
17	2238	22.3	1585	15	US-10-369-493-5570	Sequence 5570, Ap
18	2166	21.6	442	9	US-09-925-300-950	Sequence 950, App
19	1814	18.1	344	16	US-10-408-765A-1670	Sequence 1670, Ap
20	1587	15.8	793	17	US-10-777-144-3	Sequence 3, Appli
21	1587	15.8	793	17	US-10-777-145-3	Sequence 3, Appli
22	1587	15.8	793	17	US-10-777-186-3	Sequence 3, Appli
23	1565.5	15.6	802	17	US-10-777-144-1	Sequence 1, Appli
24	1565.5	15.6	802	17	US-10-777-145-1	Sequence 1, Appli
25	1565.5	15.6	802	17	US-10-777-186-1	Sequence 1, Appli
26	1565.5	15.6	807	15	US-10-264-049-3138	Sequence 3138, Ap
27	1548	15.4	1452	9	US-09-887-669-8	Sequence 8, Appli
28	1529	15.2	1452	16	US-10-408-765A-83	Sequence 83, Appl
29	1505.5	15.0	1439	9	US-09-887-669-2	Sequence 2, Appli
30	1502.5	15.0	1439	16	US-10-408-765A-284	Sequence 284, App
31	1482.5	14.8	1457	9	US-09-887-669-1	Sequence 1, Appli
32	1461	14.5	774	13	US-10-087-192-1017	Sequence 1017, Ap
33	1460.5	14.5	1444	15	US-10-058-270A-98	Sequence 98, Appl
34	1454	14.5	700	15	US-10-366-547-63	Sequence 63, Appl
35	1454	14.5	700	15	US-10-444-795B-793	Sequence 793, App
36	1454	14.5	700	16	US-10-473-127-912	Sequence 912, App
37	1454	14.5	700	16	US-10-473-127-915	Sequence 915, App
38	1454	14.5	700	16	US-10-473-127-916	Sequence 916, App
39	1454	14.5	700	16	US-10-473-127-919	Sequence 919, App
40	1454	14.5	700	16	US-10-473-127-919	Sequence 919, App
41	1454	14.5	717	16	US-10-753-267-100	Sequence 100, App
42	1454	14.5	717	16	US-10-473-127-917	Sequence 917, App
43	1454	14.5	1463	14	US-10-176-847-22	Sequence 22, Appl
44	1454	14.5	1463	14	US-10-205-823-343	Sequence 343, App
45	1444	14.4	699	15	US-10-366-547-65	Sequence 65, Appl

ALIGNMENTS

RESULT 1

- US-10-772-636-64
- Sequence 64, Application US/10772636
- Publication No. US20050042687A1
- GENERAL INFORMATION:
- APPLICANT: Kelly, Louise M.
- APPLICANT: Carroll, Joseph M.
- APPLICANT: Farlow, Deborah
- APPLICANT: Healy, Aileen
- TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
- TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 9118, 990, 17662, 81982, 630,
- TITLE OF INVENTION: 21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,
- TITLE OF INVENTION: 9465, 23544, 7366, 27417, 57259, 21844, 943, 2061, 5891,
- TITLE OF INVENTION: 9137, 13908, 14310, 17600, 25584, 27824, 28469, 38947,
- TITLE OF INVENTION: 53003, 965, 56639, 9661, 16052, 1521, 6662, 13913, 12405 OR
- TITLE OF INVENTION: 5014
- FILE REFERENCE: MPI03-015PIRNONMIM
- CURRENT APPLICATION NUMBER: US/10/772,636
- CURRENT FILING DATE: 2004-02-05
- PRIOR APPLICATION NUMBER: US 60/445,241
- PRIOR FILING DATE: 2003-02-05
- PRIOR APPLICATION NUMBER: US 60/448,389
- PRIOR FILING DATE: 2003-02-18
- PRIOR APPLICATION NUMBER: US 60/456,320
- PRIOR FILING DATE: 2003-03-20
- PRIOR APPLICATION NUMBER: US 60/460,279
- PRIOR FILING DATE: 2003-04-03
- PRIOR APPLICATION NUMBER: US 60/465,924
- PRIOR FILING DATE: 2003-04-28
- PRIOR APPLICATION NUMBER: US 60/470,052
- PRIOR FILING DATE: 2003-05-13
- PRIOR APPLICATION NUMBER: US 60/498,106
- PRIOR FILING DATE: 2003-08-26
- PRIOR APPLICATION NUMBER: US 60/500,179

; PRIOR FILING DATE: 2003-09-04									
; PRIOR APPLICATION NUMBER: US 60/502,909									
; PRIOR FILING DATE: 2003-09-15									
; PRIOR APPLICATION NUMBER: US 60/510,351									
; PRIOR FILING DATE: 2003-10-10									
; Remaining Prior Application data removed - See File Wrapper or PALM.									
; NUMBER OF SEQ ID NOS: 80									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 64									
; LENGTH: 1912									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-772-636-64									
Query Match 100.0%; Score 10042; DB 17; Length 1912;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1912; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	MVHVARLLLLLFFLRDAETPPRTTRTPVDQTVGGVASFICQATGDPRPKIWNKK	60						
Db	1	MVHVARLLLLLFFLRDAETPPRTTRTPVDQTVGGVASFICQATGDPRPKIWNKK	60						
Qy	61	GKVSQRPVEIFPDGSGSVLRIOPLRPRDEAIYECVASNNVGEISVSTRITVLREDQ	120						
Db	61	GKVSQRPVEIFPDGSGSVLRIOPLRPRDEAIYECVASNNVGEISVSTRITVLREDQ	120						
Qy	121	IPRGFPTIDMGPKVVERTRTATMLCAASGNPDPEITWPKDFLPVDTSNNNGRIKQLRS	180						
Db	121	IPRGFPTIDMGPKVVERTRTATMLCAASGNPDPEITWPKDFLPVDTSNNNGRIKQLRS	180						
Qy	181	ESIGGTPIRGALQIEQSEESDQKYECAVATNSAGTRYAPANLYVRELREVRVPPRFESI	240						
Db	181	ESIGGTPIRGALQIEQSEESDQKYECAVATNSAGTRYAPANLYVRELREVRVPPRFESI	240						
Qy	241	PPTNHEIMPQGSVNITCVAVGSMPYVYKMWLGAEADLTPEDDMPIGRNVLENDVROSANY	300						
Db	241	PPTNHEIMPQGSVNITCVAVGSMPYVYKMWLGAEADLTPEDDMPIGRNVLENDVROSANY	300						
Qy	301	TCVAMSTLGVIEAIAQITVKALPKPGTPVVTTESTATSIITLWDSGNPPEVSVYIIQHKP	360						
Db	301	TCVAMSTLGVIEAIAQITVKALPKPGTPVVTTESTATSIITLWDSGNPPEVSVYIIQHKP	360						
Qy	361	KNSEELYKEIDGVATTRYSVAGLSPYSDYEFVRVAVNNIGRGPSPBPVLTTQTSQAPSSA	420						
Db	361	KNSEELYKEIDGVATTRYSVAGLSPYSDYEFVRVAVNNIGRGPSPBPVLTTQTSQAPSSA	420						
Qy	421	PRDVQARMLSSSTLILVQWKPEPEEPNGIQGRYVYITMDPTQHVNMMKHNVDQSQTITIG	480						
Db	421	PRDVQARMLSSSTLILVQWKPEPEEPNGIQGRYVYITMDPTQHVNMMKHNVDQSQTITIG	480						
Qy	481	NLVPQKTSYVKVLAFTSIGDGPLSSDIQVITQTGVPQPLNFKAEPESSETSIILLSTPPR	540						
Db	481	NLVPQKTSYVKVLAFTSIGDGPLSSDIQVITQTGVPQPLNFKAEPESSETSIILLSTPPR	540						
Qy	541	SDTIANYELVYKDGEGHEEQRIITEPGTSYRLQGLKPNLSLYYFRLAARSPQIGASTAEI	600						
Db	541	SDTIANYELVYKDGEGHEEQRIITEPGTSYRLQGLKPNLSLYYFRLAARSPQIGASTAEI	600						
Qy	601	SARTMQSKPAPQDISCTSPSTSIILVSWQPPPEVKQNGIITEYSIKYTAVDGEDDKPH	660						
Db	601	SARTMQSKPAPQDISCTSPSTSIILVSWQPPPEVKQNGIITEYSIKYTAVDGEDDKPH	660						
Qy	661	EILGIPSDTTKYLLEQLEKWTETRIITVTAHTDVGPGPESLSVLIRTNEDVPSPGPRKVEV	720						
Db	661	EILGIPSDTTKYLLEQLEKWTETRIITVTAHTDVGPGPESLSVLIRTNEDVPSPGPRKVEV	720						
Qy	721	EAVNSTSVKVSWSRSPVFNKQHGQIRGQVHYVRMENGEPKQPMLKDVMLADAQWESFDDT	780						
Db	721	EAVNSTSVKVSWSRSPVFNKQHGQIRGQVHYVRMENGEPKQPMLKDVMLADAQWESFDDT	780						
Qy	781	TEHDMIISGLQPETSYSLSLTATYTTKGDGARSFKPLVSTTGAVPGKPRLVINHTQMTAL	840						
Db	781	TEHDMIISGLQPETSYSLSLTATYTTKGDGARSFKPLVSTTGAVPGKPRLVINHTQMTAL	840						
Qy	841	IQWHPPVDTFGLPQGYRLAFGRKDMBEPLTTLTFSEKEDHFTATD IHKGASYVFLSARNK	900						
Db	841	IQWHPPVDTFGLPQGYRLAFGRKDMBEPLTTLTFSEKEDHFTATD IHKGASYVFLSARNK	900						
Qy	901	VGFGBEMVKEISIPBEVPPTGFPQNLHSEGTSTSTSVQLSWQPPVLAERNGIITKYTYLLYRD	960						
Db	901	VGFGBEMVKEISIPBEVPPTGFPQNLHSEGTSTSTSVQLSWQPPVLAERNGIITKYTYLLYRD	960						
Qy	961	INIPLLPMEOLIVPADTTMTLGLKDPDITVDYVKVRAHTSKGPGSPSPVQFRTLPLVDQVP	1020						
Db	961	INIPLLPMEOLIVPADTTMTLGLKDPDITVDYVKVRAHTSKGPGSPSPVQFRTLPLVDQVP	1020						
Qy	1021	AKNFHVKAVMKTSVLLSWEIPENYNSAMPFKILYDDGKMWEEVDGRATOKLI VNLKPEKS	1080						
Db	1021	AKNFHVKAVMKTSVLLSWEIPENYNSAMPFKILYDDGKMWEEVDGRATOKLI VNLKPEKS	1080						
Qy	1081	YSFVLTNRNSAGGLQHRVTAKTAPDLVLRTPAFITGKTMLDGMITVQLPEVPANENIKGY	1140						
Db	1081	YSFVLTNRNSAGGLQHRVTAKTAPDLVLRTPAFITGKTMLDGMITVQLPEVPANENIKGY	1140						
Qy	1141	YIIIVPLKKSRCFKTPWESPDEMDLDELLEKISRKRRSIRYGRVVELKPYIAAHFDVLP	1200						
Db	1141	YIIIVPLKKSRCFKTPWESPDEMDLDELLEKISRKRRSIRYGRVVELKPYIAAHFDVLP	1200						
Qy	1201	TSFTLLGDDKHGFGTNNKQLQSGQEVYFVLA VMEHAESKMYATSPYSDPVVSMDLDPQPI	1260						
Db	1201	TSFTLLGDDKHGFGTNNKQLQSGQEVYFVLA VMEHAESKMYATSPYSDPVVSMDLDPQPI	1260						
Qy	1261	TDEBGLIIVWGPVLAVVFIICIVIAILLYKRAESDSRKSSI PNKKIPI SHHPTDPVE	1320						
Db	1261	TDEBGLIIVWGPVLAVVFIICIVIAILLYKRAESDSRKSSI PNKKIPI SHHPTDPVE	1320						
Qy	1321	LRRLAQTPGMAHSHPIPILELADHIERLKANDNLKFSQYESIDPGQOFTWEHSLNLYN	1380						
Db	1321	LRRLAQTPGMAHSHPIPILELADHIERLKANDNLKFSQYESIDPGQOFTWEHSLNLYN	1380						
Qy	1381	KPKNRYANVIAVDHRSVLLSAIEGIPGSDYVNAVNIYDGRYKQNAVYIATQGSLPETGDFW	1440						
Db	1381	KPKNRYANVIAVDHRSVLLSAIEGIPGSDYVNAVNIYDGRYKQNAVYIATQGSLPETGDFW	1440						
Qy	1441	RMWQORSATVVMYTKLBERSRVKCDQYWPSPRGTHGLVQVTLDTVELATYCVRTFAL	1500						
Db	1441	RMWQORSATVVMYTKLBERSRVKCDQYWPSPRGTHGLVQVTLDTVELATYCVRTFAL	1500						
Qy	1501	YKNGSSEKEEVQPOFTAMPDHGVPEHPPTFFLAFURRVKTCNPPDAGPMVWHCSAGVRT	1560						
Db	1501	YKNGSSEKEEVQPOFTAMPDHGVPEHPPTFFLAFURRVKTCNPPDAGPMVWHCSAGVRT	1560						
Qy	1561	GCFFIVIDAMLERIKHEKTVDIYGHVTLMEARQNNYMQTEBDQYIFIH DALLEAVTCGNTVE	1620						
Db	1561	GCFFIVIDAMLERIKHEKTVDIYGHVTLMEARQNNYMQTEBDQYIFIH DALLEAVTCGNTVE	1620						
Qy	1621	PARNLYAYIQKLTQIETGENVTGMELEFKRLASSKAHTSRFISANLPCKNFKRNLVINMP	1680						
Db	1621	PARNLYAYIQKLTQIETGENVTGMELEFKRLASSKAHTSRFISANLPCKNFKRNLVINMP	1680						
Qy	1681	YBSTRVCLQPIRGVSGSDYINASFIDGYRQOKAYIATQGPLAETTEDFWRMLWEHNSTIV	1740						
Db	1681	YBSTRVCLQPIRGVSGSDYINASFIDGYRQOKAYIATQGPLAETTEDFWRMLWEHNSTIV	1740						
Qy	1741	VMLTKLRENGREKCHQYWPAPERSARYQYFVVDPMAYNNPQYILREFKVTDARDGQSRTV	1800						
Db	1741	VMLTKLRENGREKCHQYWPAPERSARYQYFVVDPMAYNNPQYILREFKVTDARDGQSRTV	1800						
Qy	1801	RQFQFTDMPQGVKPSGEGFIDFIGOVHKTKEQFGODGPI SVHCSAGVORTGVFTILSIV	1860						
Db	1801	RQFQFTDMPQGVKPSGEGFIDFIGOVHKTKEQFGODGPI SVHCSAGVORTGVFTILSIV	1860						
Qy	1861	LERMRVGVVDIFQTVKMLRTORPAMVQTEDOYQFSYRAALEVGLSFDHYAT	1912						
Db	1861	LERMRVGVVDIFQTVKMLRTORPAMVQTEDOYQFSYRAALEVGLSFDHYAT	1912						

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 19:35:47 ; Search time 61 Seconds
(without alignments)
3015.843 Million cell updates/sec

Title: US-10-772-636-64
Perfect score: 10042
Sequence: 1 MVHVARLLLLLTFLRTDA.....YQFSYRAALEYLGSDHYAT 1912

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10042	100.0	1912	2 A56178	protein-tyrosine-p
2	9433.5	93.9	1894	2 C54689	protein-tyrosine-p
3	8394.5	83.6	1691	1 D54689	protein-tyrosine-p
4	7117.5	70.9	1907	2 S50893	protein-tyrosine-p
5	7032	70.0	1897	1 TDHULK	leukocyte antigen-
6	6981.5	69.5	1898	2 S46216	leukocyte antigen-
7	6971.5	69.4	1863	2 S46217	protein-tyrosine-p
8	5871.5	58.5	1499	2 I50212	protein-tyrosine-p
9	5745.5	57.2	1501	2 I58148	protein-tyrosine-p
10	5562	55.4	1496	1 A48758	protein-tyrosine-p
11	4877	48.6	2029	1 TDPELK	protein-tyrosine-p
12	4650.5	46.3	1290	2 A56493	leucocyte common a
13	4315.5	43.0	1262	1 B48758	protein-tyrosine-p
14	4127	41.1	2051	2 T30938	receptor tyrosine
15	3372	33.6	1437	2 T31093	probable protein-t
16	2996	29.8	1231	2 S53089	protein-tyrosine-p
17	2820	28.1	582	2 A57068	protein-tyrosine-p
18	2238	22.3	1585	2 T19121	probable protein-t
19	1954	19.5	398	2 I56540	protein-tyrosine-p
20	1565.5	15.6	802	1 A36065	protein-tyrosine-p
21	1553	15.5	829	1 A47373	protein-tyrosine-p
22	1542	15.4	1452	1 S17669	protein-tyrosine-p
23	1541	15.3	1452	1 S17670	protein-tyrosine-p
24	1539.5	15.3	796	1 JCL285	protein-tyrosine-p
25	1539	15.0	832	2 J68051	protein tyrosine p
26	1510	15.0	440	2 I50213	protein-tyrosine-p
27	1486	14.8	1440	2 JC6312	protein-tyrosine-p
28	1482.5	14.8	1457	1 A48066	protein-tyrosine-p
29	1455	14.5	699	2 J66132	protein-tyrosine-p

30	1454	14.5	700	1 S12053	protein-tyrosine-p
31	1372	13.7	680	2 JC8052	protein-tyrosine-p
32	1331.5	13.3	1436	2 JC5290	protein-tyrosine-p
33	1286	12.8	1442	1 B48148	protein-tyrosine-p
34	1285.5	12.8	1445	1 A48148	protein-tyrosine-p
35	1270.5	12.7	1442	2 S72441	protein-tyrosine-p
36	1266	12.6	1422	2 T42636	protein-tyrosine-p
37	1242	12.4	1301	1 A41622	protein-tyrosine-p
38	1226.5	12.2	2302	2 T14328	protein-tyrosine-p
39	1205.5	12.0	1462	1 B36182	protein-tyrosine-p
40	1204	12.0	2314	1 A46151	protein-tyrosine-p
41	1100	11.0	1273	1 TDRTLT	leukocyte common a
42	1097	10.9	818	2 T19120	hypothetical prote
43	1088	10.8	1237	2 A54080	protein-tyrosine-p
44	1060	10.6	1304	1 A46546	leukocyte common a
45	1056	10.5	1291	1 A28334	protein-tyrosine-p

ALIGNMENTS

RESULT 1

A56178
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type delta precursor - human
N:Alternate names: protein-tyrosine-phosphatase BPTP-2
C:Species: Homo sapiens (man)
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C:Accession: A56178; S12052; B4929
R:Pulido, R.; Krueger, N.X.; Serra-Pages, C.; Saito, H.; Streuli, M.
J. Biol. Chem. 270, 6722-6728, 1995
A:Title: Molecular characterization of the human transmembrane protein-tyrosine phosphatase delta isoforms.
A:Reference number: A56178; MUID:95204468; PMID:7896816
A:Accession: A56178
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1912 <PUL>
A:Cross-references: UNIPROT:P23468; GB:L38929; NID:g755652; PIDN:AAC41749.1; PID:g755653
R:Krueger, N.X.; Streuli, M.; Saito, H.
EMBO J. 9, 3241-3252, 1990
A:Title: Structural diversity and evolution of human receptor-like protein tyrosine phosphatases.
A:Reference number: S12049; MUID:91006018; PMID:2170109
A:Accession: S12052
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 350-1912 <KRU>
A:Cross-references: GB:X54133; NID:g35789; PIDN:CAA38068.1; PID:g35790
A:Note: the sequence from Fig. 5B is inconsistent with that from Fig. 5A in having 568-T
R:Adachi, M.; Sekiya, M.; Arimura, Y.; Takekawa, M.; Itoh, F.; Hinoda, Y.; Imai, K.; Yac
Cancer Res. 52, 737-740, 1992
A:Title: Protein-tyrosine phosphatase expression in pre-B cell NALM-6.
A:Reference number: A44929; MUID:92119637; PMID:1370651
A:Accession: B44929
A:Molecule type: mRNA
A:Residues: 1756-1804, 'C', 1806-1845 <ADA>
A:Cross-references: GB:S78086; NID:g243545; PIDN:AAB21147.1; PID:g243546
A:Experimental source: pre-B cell NALM-6
A:Note: sequence extracted from NCBI backbone (NCBIN:78086, NCBIP:78087)
A:Note: the authors did not report the entire codon for residue 90
C:Genetics:
A:Gene: GDB:PTPRD
A:Cross-references: GDB:I31384; OMIM:601598
A:Map position: 9p24-9p24
C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology; ogy
C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembrane
F:38-100/Domain: immunoglobulin homology <IMM1>
F:140-209/Domain: immunoglobulin homology <IMM2>
F:250-304/Domain: immunoglobulin homology <IMM3>
F:711-811/Domain: fibronectin type III repeat homology <3PR>
F:1293-1912/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F:1669-1892/Domain: protein-tyrosine-phosphatase homology <PTP2>
F:1553/Active site: Cys (phosphocysteine intermediate) #status predicted

F,1559/Binding site: substrate phosphate (Arg) #status predicted			
F,1844/Active site: Cys (phosphocysteine intermediate) #status predicted			
F,1850/Binding site: substrate phosphate (Arg) #status predicted			
Query Match 100.0%; Score 10042; DB 2; Length 1912;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1912; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MVHVARLLLLLTFPLRDAETPPRFTPTVDQTGVSGVASFICQATGDPRPKIWNKK	60
Db	1	MVHVARLLLLLTFPLRDAETPPRFTPTVDQTGVSGVASFICQATGDPRPKIWNKK	60
Qy	61	GKYSNQRFVIEFDDGSGSVLRIOPLRTPDEAIYECVANNVGEISVSTRLTVLREDQ	120
Db	61	GKYSNQRFVIEFDDGSGSVLRIOPLRTPDEAIYECVANNVGEISVSTRLTVLREDQ	120
Qy	121	IPRGFPITDMGPOLKQVBERTTATMLCAASGNPDPEITWFKDFLPDVTSSNNGRIKQLRS	180
Db	121	IPRGFPITDMGPOLKQVBERTTATMLCAASGNPDPEITWFKDFLPDVTSSNNGRIKQLRS	180
Qy	181	ESIGTPIRGALQIQSESDQGYECVATNSAGTRYSPANLVRELREVRVPFRFSI	240
Db	181	ESIGTPIRGALQIQSESDQGYECVATNSAGTRYSPANLVRELREVRVPFRFSI	240
Qy	241	PPTNHEIMPGGSVNITCVAVGSPMPYVKWMLGAEDLTPEDDMPIGRNVLNDYRQSNY	300
Db	241	PPTNHEIMPGGSVNITCVAVGSPMPYVKWMLGAEDLTPEDDMPIGRNVLNDYRQSNY	300
Qy	301	TCVAMSTLGVTEAIAQITVKALPXPPTVPTTESTATSTILTWSGPNPEPVSYIIIOHKP	360
Db	301	TCVAMSTLGVTEAIAQITVKALPXPPTVPTTESTATSTILTWSGPNPEPVSYIIIOHKP	360
Qy	361	KNSLELYKEIDGVATTRYSVAGLSYSDYERFVAVANNIGRPPSEPVLTOTSOAPSSA	420
Db	361	KNSLELYKEIDGVATTRYSVAGLSYSDYERFVAVANNIGRPPSEPVLTOTSOAPSSA	420
Qy	421	PRDVOARMLSTTILVQWKEPEENQIQGYRVYTTMDPTQHVANNMKHNVADSQIITIG	480
Db	421	PRDVOARMLSTTILVQWKEPEENQIQGYRVYTTMDPTQHVANNMKHNVADSQIITIG	480
Qy	481	NLVQKTYSVKVLAFSTIGDGPLSDIQVITQTGVPGQPLNFKAEPSESTILLSWTPPR	540
Db	481	NLVQKTYSVKVLAFSTIGDGPLSDIQVITQTGVPGQPLNFKAEPSESTILLSWTPPR	540
Qy	541	SDTTANYELVYKDGSHGEQRIITEPGTSYRLQGLKPNLSLYFRLAARSPQGLGASTAEI	600
Db	541	SDTTANYELVYKDGSHGEQRIITEPGTSYRLQGLKPNLSLYFRLAARSPQGLGASTAEI	600
Qy	601	SARTWQSKPSAPPQDISCTSPSSSTILVSWQPPPVVEKQNGIITEYSIKYTAVDGEDDKPH	660
Db	601	SARTWQSKPSAPPQDISCTSPSSSTILVSWQPPPVVEKQNGIITEYSIKYTAVDGEDDKPH	660
Qy	661	BILGIPSTTKYLLEQLBKWTEYRITVTAHTDVGPPESLSVLIRTNEDVPSGPPRKVEV	720
Db	661	BILGIPSTTKYLLEQLBKWTEYRITVTAHTDVGPPESLSVLIRTNEDVPSGPPRKVEV	720
Qy	721	EAVNSTSVKWSRSPVNKHQIRGYQVHYVRMENGEPKQPMIKVMLADAAQWEPDDT	780
Db	721	EAVNSTSVKWSRSPVNKHQIRGYQVHYVRMENGEPKQPMIKVMLADAAQWEPDDT	780
Qy	781	TEHDMIIISGLQPEYSYSLTVTAHTTTKDGARSKPLVSTTGAVPKPRLVINHTQMNTAL	840
Db	781	TEHDMIIISGLQPEYSYSLTVTAHTTTKDGARSKPLVSTTGAVPKPRLVINHTQMNTAL	840
Qy	841	IQWHPVDTFGLQYRLKFGKQMEPLTTTLEFSEKEDHFTATDIHGKASVVFRLSARNK	900
Db	841	IQWHPVDTFGLQYRLKFGKQMEPLTTTLEFSEKEDHFTATDIHGKASVVFRLSARNK	900
Qy	901	VGFGEEMVKEISIPREVTPGPQNLSHSEGTSTSVQLSQWQPVLAERNGIITKTYLLYRD	960
Db	901	VGFGEEMVKEISIPREVTPGPQNLSHSEGTSTSVQLSQWQPVLAERNGIITKTYLLYRD	960
Qy	961	INIPLLPMEQLIVPADTTMTLTGLKPDYTDYVKVRAHTSKGPGYSPSVQFRTLFPDQVF	1020

Db	961	INIPLLPMEQLIVPADTTMTLTGLKPDYTDYVKVRAHTSKGPGYSPSVQFRTLFPDQVF	1020
Qy	1021	AKNFHVKAVMKTSVLLSWEIPENYNSAMPFKILYDDGKMEVEVDGRATQKLIIVNLKEPKS	1080
Db	1021	AKNFHVKAVMKTSVLLSWEIPENYNSAMPFKILYDDGKMEVEVDGRATQKLIIVNLKEPKS	1080
Qy	1081	YSFVLNTRNGNSAGLQHRVTAKTAPDVLRTKPAFIGIKTNLDGMITVQLPEVPANENIKGY	1140
Db	1081	YSFVLNTRNGNSAGLQHRVTAKTAPDVLRTKPAFIGIKTNLDGMITVQLPEVPANENIKGY	1140
Qy	1141	YIIIVPLKKSGRGFIKPWESPDEWELDELLKEISRKRSIRYGREVELKPIYIAAHFVPLP	1200
Db	1141	YIIIVPLKKSGRGFIKPWESPDEWELDELLKEISRKRSIRYGREVELKPIYIAAHFVPLP	1200
Qy	1201	TEFTLGDGKHVGFTNKQLQSGQGYVFFVLAVNHEAESKMVATSPYSDPVVSMOLDPQPI	1260
Db	1201	TEFTLGDGKHVGFTNKQLQSGQGYVFFVLAVNHEAESKMVATSPYSDPVVSMOLDPQPI	1260
Qy	1261	TDSEEGLIWVGVPLAVVFIICIVIAIILYKRAESDSRSKSSIPNNKEIPSHHPTDPVE	1320
Db	1261	TDSEEGLIWVGVPLAVVFIICIVIAIILYKRAESDSRSKSSIPNNKEIPSHHPTDPVE	1320
Qy	1321	LRRLNFOTPGMAHPPPIPILELADHIERLKANDNLKFSQBYESIDPQOQTWEHNSLEVN	1380
Db	1321	LRRLNFOTPGMAHPPPIPILELADHIERLKANDNLKFSQBYESIDPQOQTWEHNSLEVN	1380
Qy	1381	KPKNRVANVTAYDHSRVLLSAIEGIPGSDVNNANYIDGYRKQNAVATATQGSLETGDFW	1440
Db	1381	KPKNRVANVTAYDHSRVLLSAIEGIPGSDVNNANYIDGYRKQNAVATATQGSLETGDFW	1440
Qy	1441	RMIEWORSATVVMWTKLEERSVKDQYWPISRGTTGHLVQVTLTLDVELATYCVRTFAL	1500
Db	1441	RMIEWORSATVVMWTKLEERSVKDQYWPISRGTTGHLVQVTLTLDVELATYCVRTFAL	1500
Qy	1501	YKNGSSEKREVROQFTAMPDHGVPBHPPTFLAFLRVKTCNPPDAGPMVHVCSAGVGR	1560
Db	1501	YKNGSSEKREVROQFTAMPDHGVPBHPPTFLAFLRVKTCNPPDAGPMVHVCSAGVGR	1560
Qy	1561	GCFTVIDAMLERIKHEKTVDIYGHVTLMRAQNVMTQEDQYFIHIDALLEAVTCGNTEV	1620
Db	1561	GCFTVIDAMLERIKHEKTVDIYGHVTLMRAQNVMTQEDQYFIHIDALLEAVTCGNTEV	1620
Qy	1621	PARNLYAYIOKLQIETGENVTGMELEFKELASSKAHTSRFISANLPCNPKFQRLVNIMP	1680
Db	1621	PARNLYAYIOKLQIETGENVTGMELEFKELASSKAHTSRFISANLPCNPKFQRLVNIMP	1680
Qy	1681	YESTRVCLQPIRGVEGSDYINAFIDGYROKAYIATQGPLAETTEDFWMLWEHNSTIV	1740
Db	1681	YESTRVCLQPIRGVEGSDYINAFIDGYROKAYIATQGPLAETTEDFWMLWEHNSTIV	1740
Qy	1741	VMLTKLREMGREKCHOYMPAERSARYQYFVVDPAEYVNMPOYILREFKVTDDARQGSRTV	1800
Db	1741	VMLTKLREMGREKCHOYMPAERSARYQYFVVDPAEYVNMPOYILREFKVTDDARQGSRTV	1800
Qy	1801	RQFQFTDWPQGVKPSGEGFIDFQVHKTEQFGQDGPISVHCSAGVGRGTGVTITLSIV	1860
Db	1801	RQFQFTDWPQGVKPSGEGFIDFQVHKTEQFGQDGPISVHCSAGVGRGTGVTITLSIV	1860
Qy	1861	LERMRVEGVVDIFQTVKMLRTQRPAMVQTDQYOFYSRAALEYILGSPDHVAT 1912	
Db	1861	LERMRVEGVVDIFQTVKMLRTQRPAMVQTDQYOFYSRAALEYILGSPDHVAT 1912	

RESULT 2

C54689
protein-tyrosine-phosphatase (SC 3.1.3.48), receptor type delta, splice form B precursor
N;Alternate names: MPTP delta type B/C
N;Contains: protein tyrosine phosphatase, receptor type delta, splice form C
C;Species: Mus musculus (house mouse)
C;Date: 25-Apr-1995 #sequence_revision 19-May-1995 #text_change 09-Jul-2004
C;Accession: C54689; B54689
R;Mizuno, K.; Hasegawa, K.; Katagiri, T.; Ogimoto, M.; Ichikawa, T.; Yakura, H.

Mol. Cell. Biol. 13, 5513-5523, 1993
A>Title: MPTP delta, a putative murine homolog of HPTP delta, is expressed in specialized
A'Reference number: A54689; MUID:93360986; PMID:8355697
A'Accession: C54689
A>Status: preliminary
A'Molecule type: mRNA
A'Residuals: 1-1894 <MIZ>
A'CROSS-references: UNIPROT:Q64487
A'Experimental source: brain; splice form B
A>Note: sequence inconsistent with nucleotide translation
A>Note: sequence extracted from NCBI backbone (NCBIN:137486, NCBIP:137487)
A'Accession: B54689
A>Status: preliminary
A'Molecule type: mRNA
A'Residuals: 1-352, 'H', 354-535, 'S', 537-601, 1002-1894 <MIZ>
A'Experimental source: brain; splice form C
A>Note: sequence inconsistent with nucleotide translation
A>Note: sequence extracted from NCBI backbone (NCBIN:136527, NCBIP:136530)
C'Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
C'Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester hyd
F:45-107/Domain: immunoglobulin homology <IMM1>
F:245-299/Domain: immunoglobulin homology <IMM2>
F:317-399/Domain: fibronectin type III repeat homology <FN3A>
F:1278-1894/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F:1652-1874/Domain: protein-tyrosine-phosphatase homology <PTP2>
F:1536/Active site: Cys (phosphotyrosine intermediate) #status predicted
F:1542/Binding site: substrate phosphate (Arg) #status predicted
F:1826/Active site: Cys (phosphotyrosine intermediate) #status predicted
F:1832/Binding site: substrate phosphate (Arg) #status predicted

Query Match 93.9%; Score 9433.5; DB 2; Length 1894;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1820; Conservative 28; Mismatches 37; Indels 29; Gaps 11;

Qy	1	MVHARLLLLTFLRLDAPPTPTPTDQVSGVASFICQATGDRPKLVNKK	60
Db	8	MVVRRLSLLTFLCACATPTPTPTDQVSGVASFICQATGDRPKLVNKK	67
Qy	61	GKVSQNRQFVIEFDGSGVLRIOPLRTPDEAIECVASNNVGEISVSTRLTVLREDQ	120
Db	68	GKVSQNRQFVIEFDGSGVLRIOPLRTPDEAIECVASNNVGEISVSTRLTVLREDQ	127
Qy	121	IPRGPTIDMGQPKVTRTATMLCAASGNPDPEITWFKDPLVDTSNNGRIKQLRS	180
Db	128	IPRGPTIDMGQPKVTRTATMLCAASGNPDPEITWFKDPLVDTSNNGRIKQLRS	187
Qy	181	ESIGTPIRGALQIQESDQKYEVCVATNSAGTRYAPANLYVRELREVRVPPRESI	240
Db	188	ESI-----GALQIQESDQKYEVCVATNSAGTRYAPANLYVRELREVRVPPRESI	235
Qy	241	PPTNHEIMPGGSVNITCVAVGSPMPYVKKMLGAEDLTPEDDMPIGRNVLELNDVRQSANY	300
Db	236	PPTNHEIMPGGSVNITCVAVGSPMPYVKKMLGAEDLTPEDDMPIGRNVLELNDVRQSANY	295
Qy	301	TCVAMSTLGVTEAIAQITVKALPRPGTPVTESTATSITLTWDSGNPEPVSYYIIQKP	360
Db	296	TCVAMSTLGVTEAIAQITVKALPRPGTPVTESTATSITLTWDSGNPEPVSYYIIQKP	355
Qy	361	KNSBELYKEDIGVATTRYSVAGLSPSYDYEFRVAVNNIGRPPSEVLTOTSQAPSSA	420
Db	356	KNSBEYKEDIGVATTRYSVAGLSPSYDYEFRVAVNNIGRPPSEVLTOTSQAPSSA	415
Qy	421	PRDVOARMLSTTILVQWKEPEEPNGQIQGYRVYVYTMPTQHVNNMKHNVADSQITIG	480
Db	416	PRDVOARMLSTTILVQWKEPEEPNGQIQGYRVYVYTMPTQHVNNMKHNVADSQITIG	475
Qy	481	NLVPOKTVSVKVLAFSTIGDGLSSDIQVITITQTVPGQPLNFKAPESESTILLSWTPPR	540
Db	476	NLVPOKTVSVKVLAFSTIGDGLSSDIQVITITQTVPGQPLNFKAPESESTILLSWTPPR	535
Qy	541	SDTTANYELVKGDSHGGEORITIEPGTSYRLQGLKPNLSLYYFRLASPOGLGASTAEI	600

Db	536	EDTIASVELVDRDGDQGEORITIEPGTSYRLQGLKPNLSLYYFRLASPOGLGASTAEI	595
Qy	601	SARTWQSKSPAPPDICTSPSSSILVSWQPPVPEKQNGIITEYSIKYIYAVDGEDDKPH	660
Db	596	SARTWQ-KPSAPPDICTSPSSSILVSWQPPVPEKQNGIITEYSIKYIYAVDGEDDKPH	654
Qy	661	EILGIPSDTTKYLAOLEKWTYRITVTAHTDVGPGPELSVLRTNEDVSGPPRKVEV	720
Db	655	EILGNSDITKYLAOLEKWTYRITVTAHTDVGPGPELSVLRTNEDVSGPPRKVEV	714
Qy	721	EAVNATSVKSWRSPVKNQKQIRGQVHYVVRMENGEPKQPMKDVMLADAQWFEFDT	780
Db	715	EAVNATSVKSWRSPVKNQKQIRGQVHYVVRMENGEPK-SAMLKDVMLADAQ-----	767
Qy	781	TEHMIISGLQPEPYSILVTYATYTKGDGARSKPLVSTTGAVPGKPLRLVINHTMTAL	840
Db	768	---DMIISGLQPEPYSILVTYATYTKGDGARSKPLVSTTGAVPGKPLRLVINHTMTAL	824
Qy	841	IQWHPVDTEGPIQYRLKGRKDMELTTLFSEKEDHFTATDIHKGASVYFRLSARNK	900
Db	825	IQWHPVDTEGPIQYRLKGRKDMELTTLFSEKEDHFTATDIHKGASVYFRLSARNK	883
Qy	901	VGFGEEMVKEISIEEVEPTGPPQNLHSEGTSTSVQLSWQPPVLAERNGIITKVTLLYRD	960
Db	884	VGFGEEMVKEISIEEVEPTGPPQNLHSEGTSTSVQLSWQPPVLAERNGIITKVTLLYRD	943
Qy	961	INIPLLPMEOLIVPADTMTLTGLKPDYDVKVRAHSTKSGPGYSPSVQRTLPVDO-V	1019
Db	944	INVPLLPMELIVPADTMTLTGLKPDYDVKVRAHSTKSGPGYSPSVQRTLPVDOQM	1003
Qy	1020	FAKNPHKAVNKTSLVLSWEIPENYNSAMPFK-LTYDDGKAVESVDGRATOKLVNLKPE	1078
Db	1004	FAKNPHKAVNKTSLVLSWEIPENYNSAMPFK-LTYDDGKAVESVDGRATOKLVNLKPE	1063
Qy	1079	KSYFVLNTRGNSAGGLQHRVTAKTAPDLRTKPAFICKTNLDGMITVQLPEVPANENIK	1138
Db	1064	KSYFVLNTRGNSAGGLQHRVTAKTAPDLRTKPAFICKTNLDGMITVQLPEVPANENIK	1123
Qy	1139	GYIIIVPLKSKRGKFIKPWESPDMEDELLELKEISRKRSIRYGREVELKPYIAAHFDV	1198
Db	1124	GYIIIVPLKSKRGKFIKPWESPDMEDELLELKEISRKRSIRYGREVELKPYIAAHFDV	1183
Qy	1199	LPTEFTIGDDKHGGFTNKQLQSQOEYVFFVLAHMAESKMYATSPYSDPVWSMDLPQ	1258
Db	1184	LPTEFTIGDDKHGGFTNKQLQSQOEYVFFVLAHMAESKMYATSPYSDPVWSMDLPQ	1243
Qy	1259	PITDEEGLIWWGVLAVVFIICIVAILLYKRAESDSRKSIPNNKEIPSHHPTDP	1318
Db	1244	PITDEEGLIWWGVLAVVFIICIVAILLYKRAESDSRKSIPNNKEIPSHHPTDP	1303
Qy	1319	VELARLNFQTPGMASHPEPIPIELADHIERLKANDNLKFSQEYESIDPGQOFTWEHNSLE	1378
Db	1304	VELARLNFQTPGMASHPEPIPIELADHIERLKANDNLKFSQEYESIDPGQOFTWEHNSLE	1363
Qy	1379	VNPKRYANVIAIDHSRVLLSAIEGIPGSDYNNYVINDGKRAESKSSLSKNSKEVPSHHPTDP	1438
Db	1364	VNPKRYANVIAIDHSRVLLSAIEGIPGSDYNNYVINDGKRAESKSSLSKNSKEVPSHHPTDP	1423
Qy	1439	FWRMWQRATVMMTKLEERSRVKCDQVWPSRGTEHGLVQVTLDTVELATYCVRTF	1498
Db	1424	FWRMWQRATVMMTKLEERSRVKCDQVWPSRGTEHGLVQVTLDTVELATYCVRTF	1481
Qy	1499	ALYKNGSSEKVRQFOFTAWPDHGVPEHPTPLAFLLRRVKTCTNPPDAGPMVHCSAGVG	1558
Db	1482	ALYKNGSSEKVRQFOFTAWPDHGVPEHPTPLAFLLRRVKTCTNPPDAGPMVHCSAGVG	1541
Qy	1559	RTGCFIVIDAMLERIKHEKTVDIYGHVTLMAQRYNVQTEQYIFTHDALLEAVTCNT	1618
Db	1542	RTGCFIVIDAMLERIKHEKTVDIYGHVTLMAQRYNVQTEQYIFTHDALLEAVTCNT	1601
Qy	1619	EVARNLYAYIQKLTQIETGENTVGMELBPKRLASSKAHTSRFISANLPCNPKFNRLVNI	1678
Db	1602	EVARNLYAYIQKLTQIETGENTVGMELBPKRLASSKAHTSRFISANLPCNPKFNRLVNI	1661

Qy 1679 MPYESTRVCLOPIRGVSGDYINASFIDGVRQOKAYIATQGLAETTEDFKRLWEHNST 1738
 Db 1662 MPYESGRVCLQPIRGVSGDYINASFIDGVRQOKAYIATQGLAETTEDFKRLWEHNST 1721
 Qy 1739 IVNMLTKLRMGREKCHQYWPASARSARYQVFPVDPMAEYNNPQVILREFKVTGDARDQSR 1798
 Db 1722 IVNMLTKLRMGREKCHQYWPASARSARYQVFPVDPMAEYNNPQVILREFKVTGDARD-QSR 1780
 Qy 1799 TVRQFQFTDMPQGVKSGEGFIDFQGVHKTKEQFGQDGPISVHCSAGVGTGVFTLS 1858
 Db 1781 TVRQFQFTDMPQGVKSGEGFIDFQGVHKTKEQFGQDGPISVHCSAGVGTGVFTLS 1840
 Qy 1859 IVLERMYEGVVDIFQTVKMLRTORPAMVQTEDQYQFYRAALSYLGSFQHYAT 1912
 Db 1841 IVLERMYEGVVDIFQTVKMLRTORPAMVQTEDQYQFYRAALSYLGSFQHYAT 1894

RESULT 3
 D54689
 protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type delta, splice form D precursor
 N:Alternate names: MPTP delta type D
 N:Contains: protein tyrosine phosphatase, receptor type delta, splice form A
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: D54689; A54689
 R:MiZuno, K.; Hasegawa, K.; Katagiri, T.; Ogimoto, M.; Ichikawa, T.; Yakura, H.
 Mol. Cell. Biol. 13, 5513-5523, 1993
 A:Title: MPTP delta, a putative murine homolog of HPTP delta, is expressed in specialize
 A:Reference number: A54689; MUID:93360986; PMID:8355697
 A:Accession: D54689
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1691 <MIZ>
 A:Cross-references: UNIPROT:Q64487
 A:Experimental source: brain
 A:Note: sequence inconsistent with nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:137486, NCBIPI:136537)
 A:Accession: A54689
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-398,799-1691 <MIZ>
 A:Experimental source: brain
 A:Note: sequence inconsistent with nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:136522, NCBIPI:136524)
 C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
 ogy
 C:Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester hyd
 F:42-95/Domain: immunoglobulin homology <TM3>
 F:114-196/Domain: fibronectin type III repeat homology <FN3A>
 F:1075-1691/Domain: leukocyte common antigen cytosolic domain homology <LAC>
 F:1449-1671/Domain: protein-tyrosine-phosphatase homology <PTP2>
 F:1333/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1339/Binding site: substrate phosphate (Arg) #status predicted
 F:1623/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1629/Binding site: substrate phosphate (Arg) #status predicted

Query Match 83.6%; Score 8394.5; DB 1; Length 1691;
 Best Local Similarity 95.7%; Pred. NO. 0;
 Matches 1611; Conservative 28; Mismatches 28; Indels 17; Gaps 9;

Qy 231 VRRVPPRESIPTNHEIMPGGSVNITCVAGSPMPYVKWMLGAEDLTPEDDMPIGRNVLE 290
 Db 23 VRRVPPRESIPTNHEIMPGGSVNITCVAGSPMPYVKWMLGAEDLTPEDDMPIGRNVLE 82
 Qy 291 LNDVRQSNYTCVAMSTLGVIEAIAQITVKALPKPPGTPVVTSTSTATSITLTWDSGNPEP 350
 Db 83 LNDVRQSNYTCVAMSTLGVIEAIAQITVKALPKPPGTPVVTSTSTATSITLTWDSGNPG 142
 Qy 351 VSYIIOHKPNSEBLYKEIDGATRYSVAGLSPYSEYFRVAVNNIGRGPSEPVLT 410
 Db 143 VSYIIOHKPNSEBLYKEIDGATRYSVAGLSPYSEYFRVAVNNIGRGPSEPVLT 202

Qy 411 QTSEQAPSSAPRDVQARMLSSTIILVQWKEPEBPNGQIQGYRVYVTMDPTQHVNMMKHN 470
 Db 203 QTSEQTPSSAPRDVQARMLSSTIILVQWKEPEBPNGQIQGYRVYVTMDPTQHVNMMKHN 262
 Qy 471 VADSQITITGNLVQPKTYSVKVLAFTSISGDPGLSSDIQVITQTGVPGQPLNFKAEPSSET 530
 Db 263 VADSQITITGNLVQPKTYSVKVLAFTSISGDPGLSSDIQVITQTGVPGQPLNFKAEPSSET 322
 Qy 531 SILLSTPPRSPTIANVELVYKDGHEGEORITIEPCTSRLQGLKPNLSLYFRLAARSP 590
 Db 323 SILLSTPPRSPTIASYELVDRDQGEORITIEPCTSRLQGLKPNLSLYFRLSATSP 382
 Qy 591 QGLGASTAETASRTMOSKPSAPPQDISCTSPSSSTILVSMQPPPVKQNGIIIEYSIKYT 650
 Db 383 QGLGASTAETASRTMQ-KPSAPPQDISCTSPSSSTILVSMQPPPVKQNGIIIEYSIKYA 441
 Qy 651 AVDGEDDKPHEILGIPSDTTKYLLEQLEKWTETRTITTAHTDVGPSPESISVLIRTNEDV 710
 Db 442 AVDGEDYKPEHIIIGNSDDTTKYLLEQLEKWTETRTITTAHTDVGPWPESISVLIRTDDEV 501
 Qy 711 PSGPPRKVEAVNSTSVKVSWSVPNKHOGQIRGVQVHYVRMENGEPKQGMKXDMVL 770
 Db 502 PSGPPRKVEAVNATAVKVWSRSPNKHOGQIRGVQVHYVRMENGEPK-SAMLKXDMVL 560
 Qy 771 ADAQWEDFDDTTEHDMIIISGLQPETSYSLTVAITTKGDGARSKPKLVSTTTGAVPGKPLV 830
 Db 561 ADAQ-----DMIISGLQPETSYSLTVAITTKGDGARSKPKLVSTTTGAVPGKPLV 611
 Qy 831 INHTQNTALIQWHPPVDTPGLOGYRLKPKGRKDMPLTLFSEKEDHTATDIHKAS 890
 Db 612 INHTQNTALIQWHPPVDTFG-LQGYRLKPKGRKDMPLTLFSEKEDHTATDIHKAS 670
 Qy 891 YVFLSARNKVGFEEMVKEISIPPEVPTGFPONLHSEGTTSVOLSWOPPVLAERNGI 950
 Db 671 YVFLSARNKVGFEEMVKEISIPPEIPTGFPONLHSEGTTSVOLSWOPPVLAERNGV 730
 Qy 951 ITKYTLTYLRDINIPLLPMEQLIVPADTTMTLTGLKPDTTYDVKVRAHTSKGPPSPSVQ 1010
 Db 731 ITKYTLTYLRDINVPLLPMEHLIVPADTSMTLTGLKSDTTYDVKVRAHTSKRGPPSPSVQ 790
 Qy 1011 FRTLPLVDQ-VFARNFHVKAVMKTSVLLSWEI PENYNSAMPFK-ILYDDGKQVEVDGRAT 1068
 Db 791 FRTLPLVDQSMFAKNFHVKAVMKTSVLLSWEI PENYNPAIILSKFFMMMDGKQVEVDGRAT 850
 Qy 1069 QKLVNLPKESYFVLTVNKGNSAGGLOHEVTAKTAPDVLRTKPAFIGTKNLDGMITVOL 1128
 Db 851 QKLVNLPKESYFVLTVNKGNSAGGLOHVRTAKTAPDVLRTKPAFIGTKNLDGMITVOL 910
 Qy 1129 PEVPANENIKGYIIIVPLKKSRSKFKIPKWPESPEMELDELLKEISRKRSIRYGREVEL 1188
 Db 911 PDVPANENIKGYIIIVPLKKSRSKFKIPKWPESPEMELDELLKEISRKRSIRYGREVEL 970
 Qy 1189 KPYIAAHFDVLPTFTLGDGDKHYGGFTNKLOQSQGEYVFFVLAVMEHAESKMYATSPYS 1248
 Db 971 KPYIAAHFDVLPTFTLGDGDKHYGGFTNKLOQSQGEYVFFVLAVMEHAESKMYATSPYS 1030
 Qy 1249 PVVSMDLDPQITDEERGLIWWVGPVLAVFIICIVIAILLYKRAESDRKSISPNKK 1308
 Db 1031 PVVSMDLDPQITDEERGLIWWVGPVLAVFIICIVIAILLYKRAESDRKSISLPSNK 1090
 Qy 1309 EIPSHHPTDPEVLRRLNFOTPGMAASHPPPIPILEADHIERLKANDNLKFSQYESIDPGQ 1368
 Db 1091 EVPSHHPTDPEVLRRLNFOTPGMAASHPPPIPILEADHIERLKANDNLKFSQYESIDPGQ 1150
 Qy 1369 QFTWEHNLVKNPKRNYANVIADHSRVLLSAIEGIPGSDYNNYVINDGVRKQNYAT 1428
 Db 1151 QFTWEHNLVKNPKRNYANVIADHSRVLLSAIEGIPGSDYNNYVINDGVRKQNYAT 1210
 Qy 1429 QGSIPTETFGDFWMIWEQSRATVVMVTKLEERSVKCDQWYPSRGTEHGLVQVTLDDTV 1488
 Db 1211 QGSIPTETFGDFWMIWEQ-EATVVMVTKLEERSVKCDQWYPSRGTEHGLVQVTLDDTV 1269
 Qy 1489 ELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHPTTFLAFLRRVKTNCNPPDAGP 1548

Db 1270 EL-TYCVTFALYNGSSSEKRVQFQFAWPDHGVPEHPTFFLAFLRVKTCNPPDAGP 1328
Qy 1549 MVVHCSAGVGRGCFIVIDAMLERIKHEKTVDIYGHVTLMRQRMVQTEQDYIFIHDA 1608
Db 1329 MVVHCSAGVGRGCFIVIDAMLERIKHEKTVDIYGHVTLMRQRMVQTEQDYIFIHDA 1388
Qy 1609 LLEAVTCGNTEVPARNLVAIYQKLTQIETGENVTGMELEFKELASSKAHTSRFISANLPC 1668
Db 1389 LLEAVTCGNTEVPARNLVAIYQKLTQIETGENVTGMELEFKELASSKAHTSRFISANLPC 1448
Qy 1669 NKFNRLNINIMPEYSTRVCLQPIRGVEGSDYINAFIDGYROOKAYIATQGPLAETTEDF 1728
Db 1449 NKFNRLNINIMPEYSTRVCLQPIRGVEGSDYINAFIDGYROOKAYIATQGPLAETTEDF 1508
Qy 1729 WRMLWEHNSITVWMLTKLREMGREKCHQYPAERSARYQYFVVDPMAYNMPQYILRBFK 1788
Db 1509 WRMLWEHNSITVWMLTKLREMGREKCHQYPAERSARYQYFVVDPMAYNMPQYILRBFK 1568
Qy 1789 VTDARDGOSRTVROQFTDWEPEQGVKSGEGFIDFVGQVHKTKEQFGQDGPISVHCSAGV 1848
Db 1569 VTDARDGOSRTVROQFTDWEPEQGVKSGEGFIDFVGQVHKTKEQFGQDGPISVHCSAGV 1627
Qy 1849 GRTGVFTLSVLEMRVYEGVVDIFQTVKMLRTQRPAMVQTEQDYQFSYRAALEVILGSFD 1908
Db 1628 GRTGVFTLSVLEMRVYEGVVDIFQTVKMLRTQRPAMVQTEQDYQFSYRAALEVILGSFD 1687
Qy 1909 HYAT 1912
Db 1688 HYAT 1691

RESULT 4
S50893
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type sigma precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S50893; S40281
R:Wagner, J.; Boerboom, D.; Tremblay, M.L.
Eur. J. Biochem. 226, 773-782, 1994
A:Title: Molecular cloning and tissue-specific RNA processing of a murine receptor-type
A:Reference number: S50893; MUID:95112841; PMID:7529177
A:Accession: S50893
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1907 <WAG>
A:Cross-references: UNIPROT:Q64494; EMBL:X82288; NID:G587483; PIDN:CAA57732.1; PID:G5874
R:Hendriks, W.; Brugman, C.; Zeeuwen, P.; Schepens, J.; Wieringa, B.
submitted to the EMBL Data Library, June 1993
A:Description: Assessment of the expression levels of murine protein-tyrosine phosphatase
A:Reference number: S40280
A:Accession: S40281
A:Molecule type: mRNA
A:Residues: 1441-1501, 'E', 1503-1546 <HEN>
A:Cross-references: EMBL:Z23050; NID:G438137; PIDN:CAA80585.1; PID:G438138
C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
cgy
C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembrane
F:149-209/Domain: immunoglobulin homology <IMW1>
F:246-300/Domain: immunoglobulin homology <IMW2>
F:413-506/Domain: fibronectin type III repeat homology <3PR>
F:1188-1907/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F:1375-1596/Domain: protein-tyrosine-phosphatase homology <PTR1>
F:1664-1887/Domain: protein-tyrosine-phosphatase homology <PTR2>
F:1548/Active site: Cys (phosphotyrosine intermediate) #status predicted
F:1554/Binding site: substrate phosphate (Arg) #status predicted
F:1839/Active site: Cys (phosphotyrosine intermediate) #status predicted
F:1845/Binding site: substrate phosphate (Arg) #status predicted

Query Match 70.9%; Score 7117.5; DB 2; Length 1907;
Best Local Similarity 70.3%; Pred. No. 0;
Matches 1351; Conservative 234; Mismatches 293; Indels 45; Gaps 16;

Qy 4 VARLLLLLTFFLTARDAETPRFTTRTPVDQTVSGGVASFICQATGDRPKLVNKKCKK 63
Db 16 VGLFLVLLARCL---AEPPPRFIREPKDQIGVSGGVASFVQCATGDRPKRVWNKKCKK 72
Qy 64 VSNRFEVIEFDGSGSVLRIQPLRTPRDEAIYECVSNVYGEISVSTRLTVLRDQIPR 123
Db 73 VNSQRFETIDDESSGAVLRIQPLRTPRDENVYECVAGNSVGEITIAKLTVLREDQLPP 132
Qy 124 GFPIIDMGPKQLKVERTRTATMLCAASGNPDPEITWFKDFLPVDTSSNNNGRIKQLRSSSI 183
Db 133 GFPIIDMGPKQLKVERTRTATMLCAASGNPDPEITWFKDFLPVDPSPASNGRIKQLRS--- 189
Qy 184 GGTPIRGALQLEQSESDQGGKYECVATNSAGTRYSAPANLVVRELRRVRRVPPRPSIPT 243
Db 190 -----GALQIESSEETDQGGKYECVATNSAGTRYSAPANLVVRRVRRVPPRPSIPT 239
Qy 244 NHEIMPGGSVNITCVAVGSPMPYVKMMLGAEDLTPEDDMPIGRNVLELNDVRSANYTCV 303
Db 240 SHEIMPGGSVNITCVAVGSPMPYVKMMLGAEDLTPEDDMPVGRNVLELTDVKDSANYTCV 299
Qy 304 AMSTLGVIBIAQITVKALPKPOTPVVTSTATSIITLWDSGNPEPVSYIIOHKPKNS 363
Db 300 AMSSLGIVIEVAQITVKSLPKAPGTPVVTENTATSIITWDSGNPDVSYVIVSVKSKSQ 359
Qy 364 BELYKEIDGVATTSYVAGLSPYSDYEFVRVAVNNIGRGPSEPVLTQTSQAPSSAPRD 423
Db 360 DGPVQIKEDITTTTYSIGGLSPNSEYIWSAVNSIGQGPSESVVTRTGSQAPASAPRN 419
Qy 424 VQARMLSTTILVQWKEPEBPNGOIQGVYVYVMDPTQHVNNMMKNVADSOITIGNLV 483
Db 420 VQARMLSATMIQWKEPEBPNGOIQGVYVYVMDPTQHVNNMMKNVADSOITIGNLV 479
Qy 484 PQKTSYVKVLAFTSIGDPLSSDIQVITQTGVCPQPLNFKAPSESETSIILSWTPPRSDT 543
Db 480 EDETYTVRLAFTSVGDPLSDPIQVKTQGVCPQPMNLRAEAKSETSIGLSWAPRQES 539
Qy 544 IANYELVYKDGEGEEORIIEPCTSYRLQGLKNSLYYFRLAARSQGLGASTAEISAR 603
Db 540 VIKYELLFREGDRGREGVTRFDPTTAFVVEDLKNTEYAFRLAARSQGLGASTAEISAR 599
Qy 604 TMOGKSPAPQDISCTSPSTSIILVSWQPPPEKQNGIITEYSIKYTAVDGEDDKPHEIL 663
Db 600 TLQAKSPAPQDVKCTSLRSTAILVSWRPPPEPHNGALVGYSVRYRPLGSEDDPKKEVN 659
Qy 664 GIPSDTTKYLEQLEKWEYRITVTAHTDVGPGPESLSVLRINEDVSPGPPRKYVEAV 723
Db 660 NIPPTTQILLALEKWEYRITVTAHTDVGPGPESLSVLRINEDVSPGPPRKYVEAV 719
Qy 724 NSTSVKYSWSPVNPKNQGOIRGVQVHYVRMENGEPKQPMKDVMLADQWEDDTTEH 783
Db 720 NATAIRVLWSPTPGRQHQIRGVQVHYVRMENGEPKQPMKDVMLADQWEDDTTEH 770
Qy 784 DMIISGLQPTSYSLTVATVYTKDGAARSKPLVSTTGAVPKPLRVNTHOMNTALQW 843
Db 771 EMVITNLOPFAISITVAATVMKDGARSKPQVVTGKAVLGRPTLSVQQTPEGSLARW 830
Qy 844 HPPVDTF-GPLQGYRLKFKGKMEPLTTLFSEKEDHFTATDIHKGASYYFRLSARNKVG 902
Db 831 EPPGDAADPVLGYRLQFGREDRAP-ATLELAWEREF-AAPAHKATGYVFLAARGGG 888
Qy 903 FGEWVKEISIPREVPFGFQNLHSEGTST-SVQLSWQPPVLAERNGITKYTLLYRDI 961
Db 889 LGEEAAAALSIPEADPRGFPQILGAAGNVGAGSVLLRMLPVPAPRNGAIKYTVSVREA 948
Qy 962 NIPLLPMEQLIVP-----ADTTMTLGLKPDITTVKVRHAHTSGPGYSPSVQFRTLVP 1016
Db 949 GAGPATETELAAAPGASTALTGLRLPETGYELRVRAHTRGPGFPFPLIYR-LAR 1007
Qy 1017 DQVFAKNFHVKAVMTKSVLLSWEPINENYNSAMPFKILYDDGKAVVEYDGRATOKLIYNLK 1076
Db 1008 DPVSPKPKFKVMIMKTSVLLSWEPFDPNYSPTPKIYQNG--LTLDVDRGTTKKLIITHLK 1065
Qy 1077 PEKSYSVLTNRGNSAGGLQHRVTAKTADPVLRTKPAFIKTNLDGMITVOLPEVPANEN 1136

Db	1066	PHTYFNVLNRRSSGLGLOQTVTARTAFNMLSKSPVAPKPDNDGFIIVVLPDQGPVPT	1125
Qy	1137	IKGYIIIVPLKSR-GKFIKPWSPDEMDLLEKISR-KRRSIRYGREVEL-KPYIA	1193
Db	1126	VONYFIWVPLKSRGQFPVLLGSPEDMDLEELIQDISRLQORQVHRHSQLEVPFYIA	1185
Qy	1194	AHFDVLPTEFTLGDGKHGGFTNKQLSGQGYVFFVLVAWEHAKMYATSPYSDPVVSM	1253
Db	1186	ARFSILPAVHPGNQKQVGGFNDRGLEPHGRHYVFLVLAQLQNE-PTFAASPFSDPFQLD	1244
Qy	1254	DLDPQITDEEGLIWWVGPVLAFFIICIVIAILLYKR----KRAESDSKSKSIPNNKE	1309
Db	1245	NPDPQPIVDGEGLIWIGPVLAVFFIICIVIAILLYKRPDSRKSEPTKCLLNAD	1304
Qy	1310	IPSHPTDPVBLRLNFPQMGASHPTPIPILELADHIERLKANLNKFSQYESIDPGQQ	1369
Db	1305	LAPHPKDPVEMRRINFQTPGMLSHPTPIPTDMAEHMERLKANDSLKSQYESIDPGQQ	1364
Qy	1370	FTWEHSLNLEVNKPKRYANVTAYDHSRVLLSAISGIPGSDYVNAVYIDYRKONAYATQ	1429
Db	1365	FTWEHSLNLEANKPKRYANVTAYDHSRVILQPLEGIMGSDYINANYVDYRRQWAYATQ	1424
Qy	1430	GSLPETGDFWRMIWEQRSATVWMTKLEERSRVKCDQWPSRGTTGHLVQVTLTDTVE	1489
Db	1425	GPLPETGDFWRVWVWEQRSATVWMTLREBSRKICDQWPNRGTTGYFIQVTLTDTME	1484
Qy	1490	LATCFVRTFALYKNGSSKREVRQFTAWPDHGVPEHTPFLAFLRVKTCNPPDAGPM	1549
Db	1485	LATCFVRTFSLKNGSSKREVRHFQFTAWPDHGVPEYPTPFLAFLRVKTCNPPDAGPI	1544
Qy	1550	VVHCSAGVTRGCTFIVIDAMLERIKHEKTVDIYGHVTLRMAORNVMQTEDOYFIHDAL	1609
Db	1545	VVHCSAGVTRGCTFIVIDAMLERIKTEKTVDVYGHVTLRMSQRNVMQTEDOYGFIEAL	1604
Qy	1610	LEAVTCGNTEVPARNLVAYIOKLTQIETGENVNTGMELEFKRLASSKAHTSRFISANLPCN	1669
Db	1605	LEAVSCGNTEVPARSLTYIQKLAQVEGHEVNTGMELEFKRLASSKAHTSRFITASLPCN	1664
Qy	1670	KFKNRLNIMPYESTRVCQLPIRGVSGSDYINASFIDGYROOKAYIATQGPLAETTEDFW	1729
Db	1665	KFKNRLNILPYESSRVCQLPIRGVSGSDYINASFIDGYROOKAYIATQGPLAETTEDFW	1724
Qy	1730	RMLHEHNSIIVMLTKREMREKCHOYPAERSARYQYFVVDMAEYNNMPOYILREPKV	1789
Db	1725	RALMENNSTIIVMLTKREMREKCHOYPAERSARYQYFVVDMAEYNNMPOYILREPKV	1784
Qy	1790	TDARDGQSRTVRQFOFTDWPQGVPKSGEGFIDFTGVHKTKEQFGDGPISVHCSAGVG	1849
Db	1785	TDARDGQSRTVRQFOFTDWPQGVPKSGEGFIDFTGVHKTKEQFGDGPISVHCSAGVG	1844
Qy	1850	RTGVFTILSVILERNRYEGVDIFQTVKMLRTQRPAMVQTEDOYFVSRAALEYLGSPDH	1909
Db	1845	RTGVFTILSVILERNRYEGVDIFQTVKMLRTQRPAMVQTEDEYQFCFQAALEYLGSPDH	1904
Qy	1910	YAT 1912	
Db	1905	YAT 1907	

RESULT 5
TDHULK
leukocyte antigen-related protein precursor - human
N:Alternate names: leukocyte common antigen homolog
N:Contains: protein-tyrosine-phosphatase (EC 3.1.3.48)
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: S03841; J00051
R:Streuli, M.; Krueger, N.X.; Hall, L.R.; Schloesman, S.F.; Saito, H.
J. Exp. Med. 168, 1523-1530, 1988
A:Title: A new member of the immunoglobulin superfamily that has a cytoplasmic region hc
A:Reference number: J00051; MUID:89035978; PMID:2972792
A:Accession: S03841

A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1897 <STR>
A:Cross-references: UNIPROT:P10586; EMBL:Y00815; NID:g34266; PIDN:CAA68754.1; PID:g34267
C:Genetics:
A:Gene: GDB:PTPRF; LAR
A:Cross-references: GDB:120138; OMIM:179590
A:Map Position: lp34-lp34
C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
ogy
C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembrane
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1897/Product: leukocyte antigen-related protein #status predicted <MAT>
F:17-1250/Domain: extracellular #status predicted <EXT>
F:37-99/Domain: immunoglobulin homology <IMM1>
F:139-199/Domain: immunoglobulin homology <IMM2>
F:236-290/Domain: immunoglobulin homology <IMM3>
F:308-390/Domain: fibronectin type III repeat homology <FN3A>
F:403-489/Domain: fibronectin type III repeat homology <FN3B>
F:501-583/Domain: fibronectin type III repeat homology <FN3C>
F:586-585/Domain: fibronectin type III repeat homology <FN3D>
F:698-798/Domain: fibronectin type III repeat homology <FN3E>
F:810-893/Domain: fibronectin type III repeat homology <FN3F>
F:905-989/Domain: fibronectin type III repeat homology <FN3G>
F:1001-1078/Domain: fibronectin type III repeat homology <FN3H>
F:1251-1274/Domain: transmembrane #status predicted <TM>
F:1275-1897/Domain: intracellular #status predicted <INT>
F:1285-1897/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F:1365-1586/Domain: protein-tyrosine-phosphatase homology <PTP1>
F:1654-1877/Domain: protein-tyrosine-phosphatase homology <PTP2>
F:44-97,146-197,243-288/Diulfide bonds: #status predicted
F:107,240,285,711,956/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:1338/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1544/Binding site: substrate phosphate (Arg) #status predicted
F:1829/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1835/Binding site: substrate phosphate (Arg) #status predicted

Query Match	70.0%	Score	7032;	DB	1;	Length	1897;
Best Local Similarity	69.8%	Pred. No.	0;				
Matches	1338;	Conservative	235;	Mismatches	305;	Indels	38;
Gaps	14;						
Qy	9	LLLLTFELRTDAETPPRFTPTVDQTVGSGVASFICQATGDRPKIVMKNKKGVSNQR	68				
Db	8	LVMLGLVAGAGSDSKPVFIKVPEDQTVGSGVASFVQATGPKPRITMKNKKGVSNQR	67				
Qy	69	FEVIEFDGSGVLRIQPLRPDEALIEYCVASNNVCEISVSTLTVLREDQIPRGPTI	128				
Db	68	FEVIEFDGSGVLRIQPLRVQDEALIEYCVATNSLGEINTSAKSLVLEBEEQLPPGFSI	127				
Qy	129	DMGPQLKVVTRTATMLCAASGNPDPEITWFKDFLPVDTNNNGRIKQLRSSEIGGTP	188				
Db	128	DMGPQLKVVTRTATMLCAAGNPDPEISNFKDFLPVDPATNSGRIRKQLRS	179				
Qy	189	RGALQIEQSEESDQKYEVCVATNSAGTRYAPANLYVRELREVRVPRPSIPTNHEIM	248				
Db	180	-GALQIESEESDQKYEVCVATNSAGTRYAPANLYVRLVAVRSEAVRARTGEQAPSPRRVQARM	234				
Qy	249	PGGSVNTCTVAVGSPMPYKVMGLCAEDLTPEDDMPICGRNVLELNDVQSNYTCVAMSTL	308				
Db	235	PGGSVNTCTVAVGAPMPYKVMGMGAELTKEDEMPVGRNVLELNSVRSANYTCVAILSSL	294				
Qy	309	GVTEAIAQITVKALPKPPGTPVVTTESTATSTLTWDSGNPEPVSYYIIQHKPKNSEELYK	368				
Db	295	GMTEAIAQITVKALPKPPIDLVTTTATSTLTWDSGNSEPVYYIIQVRAAGTEGPFQ	354				
Qy	369	EIDGVATTRYSVAGLSYSDYEFVRVAVNNIGRPPSEPVLTQTSQAPSPAPSDVQARM	428				
Db	355	EVDGVATTRYSIGLSYSPFEYAFRLAVNSIGRPPSEAVRARTGEQAPSPAPSDVQARM	414				
Qy	429	LSSTTILVQWKEPEEPNGQIQGVYVYTMPTQHVNNMKNHGVADSOITITGNLVPOKTY	488				
Db	415	LSASTMLVQWKEPEEPNGLVRGVYVYTPDSRRRPFNAWHKHNTDAGLLTIVGSLPGITY	474				